TIGHT BINDING BOOK

Pages are missing

INTERMEDIATE
ELECTRICAL THEORY

BY

H. W. HECKSTALL-SMITH, M.A.

REVISED EDITION

WITH COLOURED CHART SHOWING THE THREE NATURAL RADIOACTIVE SERIES, THE NEPTUNIUM SERIES, AND SIX ARTIFICIAL SUBSIDIARY SERIES

J. M. DENT AND SONS LTD.
BEDFORD ST. LONDON W.C.2

by
J. M. DENT & SONS LTD
Aldine House · Bedford Street · London
Made in Great Britain
at
The Aldine Press · Letchworth · Herts
First published 1932
Last revised 1955

INTRODUCTION

Καὶ δὴ καὶ τὰ τῶν ὑδάτων πάντα ῥεύματα, ἔτι δὲ τὰ τῶν κεραυνῶν πτώματα καὶ τὰ θαυμαζόμενα ἠλέκτρων περὶ τῆς ἕλξεως καὶ τῶν Ἡρακλείων λίθων, πάντων τούτων ὁλκὴ μὲν οὐκ ἔστιν οὐδενί ποτε, τὸ δὲ κενὸν εἶναι μηδὲν περιωθεῖν τε αὑτὰ ταῦτα εἰς ἄλληλα, τό τε διακρινόμενα καὶ συγκρινόμενα πρὸς τὴν αὑτῶν διαμειβόμενα ἕδραν ἕκαστα ἰέναι πάντα, τούτοις τοῖς παθήμασι πρὸς ἄλληλα συμπλεχθεῖσι τεθαυματουργημένα τῷ κατὰ τρόπον ζητοῦντι φανήσεται.

Furthermore, as regards all flowings of waters, and fallings of thunderbolts, and the marvels concerning the attraction of amber and of the Heraclean stone—not one of these ever possesses any real power of attraction; but the fact that there is no void, and that these bodies propel themselves round, one into another, and that according as they separate or unite they all exchange places and proceed severally each to its own region—it is by means of these complex and reciprocal processes that such marvels are wrought, as will be evident to him who investigates them properly.

PLATO, *Timaeus*, 80 C. (Loeb Translation,
by kind permission of the Editors.)

PREFACE

THIS book deals with the classical and modern electrical theory required from the beginning of sixth-form work at schools to the end of the introductory courses at universities.

Part I contains no Calculus at all, and covers the whole of the classical part of the subject as far as is possible with this limitation.

Part II uses Calculus freely, and gives much of the fundamental mathematical theory for such an examination as Part I of the Natural Science Tripos at Cambridge. The mathematical foundations of Electrostatics are given in one short chapter; of Electromagnetism in two very short chapters, each complete in itself, but adopting different methods; of Electromagnetic Induction, Magnetic Materials, and Alternating Currents in one chapter each.

The scheme of having a minimum of mathematics in Part I, and a minimum of description in Part II, is the result of teaching experience over a long period.

I have found that competent mathematical students have great difficulty in grasping the mathematical parts of Physics because the mathematics are usually embedded in descriptive matter.

Weak mathematicians, on the other hand, seem appalled by the mathematics they meet, and givè way to despair over physical discussions which they are really quite capable of grasping.

In Part II, Chapter II and Chapter III derive the same fundamental results of Electromagnetism. Chapter II uses Ampère's Magnetic Shell method, and Chapter III the Current-element method. It seems well to realize that these two methods, though equally important and leading to identical results, are quite distinct. The former is better for problems involving Magnetic Potential; the latter for problems involving forces; and either may be used for problems about Magnetic Intensity.

Though the book begins from first principles, and could be used by a student who had never tackled the subject at all, I had in mind one who had reached matriculation standard in Electricity and Magnetism.

I wrote Part III after examining the content of the courses of introductory lectures being given in modern physics at several universities in 1948. As such lectures must be abreast of developments from year to year, it would be absurd to try to cover the same ground in a text-book. Instead I have tried to provide a background, a coherent account of what a lecturer would like to feel his audience had read and understood.

In attacking this difficult task I have experienced so much pleasure that I can only hope a little of what I felt may be communicated to my readers in spite of the cares and burdens of work for a degree course in Physics during the rush-hour.

True learning cannot advance without true delight in learning. Every one who has done serious work in science knows that what draws him on is not duty but joy; and in writing Part III, as twenty years earlier with much of Part II and some of Part I, I certainly did not miss my share.

H. W. HECKSTALL-SMITH.

1955 EDITION

THE final chapter, 'Particles,' has been revised to include discoveries since 1949; the section on reference books, pp. 668–671, has been revised; and a long note has been added to the back of the chart of the four radioactive series.

There is a new Note on Metre-Kilogramme-Second Units, written because electrical engineering students have been finding useful both the fundamental electromagnetic theory in Part II and the particle-physics theory in Part III, now that an era of atomic production of electrical power is upon us.

This way of tackling the problem has kept the price down without losing well-tested accounts of essential theory. A new book covering the same ground would now cost more than double the price of this one. I hope the new Note will help engineers without hindering physicists.

H. W. HECKSTALL-SMITH.

ACKNOWLEDGMENTS

My thanks are due to the following for the very great help they have given me :—

Dr. E. J. Holmyard, for suggesting the work, and for his encouragement while it was in progress.

Mr. J. A. Ratcliffe, Fellow and Lecturer of Sidney Sussex College, Cambridge, for reading through the whole proofs of Part II, and for many suggestions in both Parts.

Sir J. C. Irvine, F.R.S., Principal of St. Andrews University, Mr. Clarence Smith of University College, Southampton, and Mr. B. R. Miles and Mr. M. D. Prain of Stowe, for proof-reading and suggestions in Chapters XI and XII of Part I.

Mr. R. T. Lattey and Mr. H. F. Biggs, of the Electrical Laboratory, Oxford, for suggestions in connection with the Electrostatics.

Mr. B. R. Miles and H. V. Kemp, of Stowe, for reading the proofs of Part I.

H. V. Kemp, D. L. Reeves, J. B. Sopper, J. K. Todd, and P. D. Ward of Stowe, for working out questions and providing a practice-ground for the exposition of the mathematical parts of Part II.

Mr. R. L. Knowles of Messrs. J. M. Dent and Sons for redrawing my original drawings and miraculously understanding what they were intended to represent.

Mr. Osborne of Messrs. F. E. Becker and Co. for some drawings of rather complicated apparatus.

The Cambridge Instrument Co., Messrs. F. E. Becker and Co., Messrs. W. G. Pye and Co., Messrs. Gambrell and Co., Messrs. Philip Harris and Co., Messrs. Constable and Co. and the D.P. Battery Co., of Bakewell, Derbyshire, for permission to reprint or copy illustrations from their catalogues.

The Cambridge Instrument Co. for permission to reprint several long extracts from their catalogues dealing both with

general galvanometer problems and with detailed descriptions of particular apparatus, without which it would have been difficult, if not impossible, for me to deal with Chapter VI and VII of Part II in what seemed to me a satisfactory manner.

Mr. Barron, of the Cambridge Instrument Co.'s London branch, for his help and advice, both in connection with the extracts from the Company's catalogues and in reading the proofs of these and other parts of the book.

Mr. P. G. Hunter, of Stowe, for suggesting the passage which serves as an Introduction.

I am very grateful to Dr. J. L. Hamshere, of Dartington Hall School, for his care in the work of reading the final page proofs of Part III.

1950 EDITION

I SHOULD like to express my strong sense of indebtedness to the following for their generous help:

Dr. D. H. Wilkinson of Jesus College, Cambridge, and the Cavendish Laboratory, and Mr. H. D. Griffith of the Department of Natural Philosophy, Aberdeen University, for criticizing and discussing with me in detail the whole of the galley proofs of Part III.

Dr. C. B. Childs of the Department of Natural Philosophy, Edinburgh University, for reading the galley proofs of Part III and for much constructive criticism.

Dr. Wilkinson, Mr. Griffith, Dr. Childs—and also Dr. Ashmead of the Cavendish Laboratory—for making suggestions about the content of Part III in the light of their experience in connection with the university courses they direct.

Mr. Michael Milford and Mr. D. S. Mumford of Repton School and Mr. G. H. Isitt of Christ's College, Brecon, for suggestions in connection with sixth-form courses.

Sir Charles Darwin, F.R.S., for permission to quote from his book, *The New Conceptions of Matter*, and for detailed criticism of the part of Chapter II, Part III, which deals with Bohr's theory of the one-electron atom.

Prof. C. W. Davies of the Department of Chemistry, University College, Aberystwyth, for vetting Chapters XI and XII of Part I and my corrections to them.

Mr. Griffith, Dr. Childs, Mr. Milford, Mr. Isitt, Mr. Barron of the Cambridge Instrument Co., and Mr. I. C. Jones of the

Department of Physics, University College, Aberystwyth, for advice about the revision of Parts I and II.

Prof. E. N. da C. Andrade for permission to quote from his book *The Atom and its Energy*.

Prof. C. F. Powell and Mr. G. P. S. Occhialini, and the Oxford University Press, for permission to reproduce illustrations from *Nuclear Physics in Photographs*.

The Telecommunications Research Establishments of the Ministry of Supply for their help in connection with accelerating apparatus.

The British Tungsram Radio Works Ltd. for their help in connection with triodes.

Mr. H. W. Marsh and Mr. Anthony Dent, of Messrs. J. M. Dent & Sons Ltd., for their unfailing sympathy, help, and forbearance; and Mr. J. B. Meachem for redrawing my very rough sketches so well that they are fit to serve as illustrations of the text.

1955 EDITION

I AM most grateful to Mr. H. D. Griffith, M.A., of the Department of Natural Philosophy in the University of Aberdeen, for his general advice and criticism; to Dr. W. R. C. Coode-Adams, M.A., M.Sc., Ph.D. (Cantab.), B.Sc.(Lond.), M.I.E.E., F.R.I.C., Principal of Faraday House Electrical Engineering College, for suggesting the Note on the Metre-Kilogramme-Second system of units, and for his constructive criticism; and to Mr. L. H. A. Carr, M.Sc.Tech., M.I.E.E., for considerable improvements in the final form of the Note.

H. W. HECKSTALL-SMITH.

*

CONTENTS

PART I

PART II

PART III

PART I

CHAPTER I

THE ELECTRIC CHARGE

Electric Charges—The Electroscope—Induction of Charges—Distribution of Charge on an Insulated Conductor—The Wimshurst Machine—The Electric Wind—Faraday's Ice-pail Experiment—Coulomb's Experiment—Cavendish's Experiment and the Inverse Square Law—Quantity of Charge—Definition of Unit Charge—Simple Numerical Problems—Experimental Work.

Electric Charges

It has been known since the time of Thales of Miletus, who flourished in the sixth century B.C., that a piece of amber rubbed with dry fur acquires the property of attracting light bodies such as scraps of paper. Such a piece of amber is described as "electrified" or "charged," probably because ἤλεκτρον is the Greek for amber.

Dr. William Gilbert (1540–1603), Queen Elizabeth's physician, repeated and amplified the electrical experiments which had been begun more than two thousand years before his time, and by the end of the seventeenth century much was known about the behaviour of electrical charges.

It was soon found that glass rubbed with silk also acquired the property of attracting light objects.

Gilbert's experiments also showed that two pieces of charged amber would repel one another, that two pieces of charged glass would repel one another, and that charged amber and charged glass would attract one another.

When only the charging of amber was known, it was supposed that a special substance, called "electricity," appeared on the amber when it was rubbed. So the strict original meaning of the word "electricity" is thus, "whatever gives amber the power of attracting light objects." This meaning is now of historical interest only.

When the electrification of glass was discovered it was

supposed that two kinds of electricity existed, and that like kinds repelled each other and unlike kinds attracted each other. The charge on glass rubbed with silk was called "vitreous," and that on amber rubbed with fur was called "resinous." This theory was soon found to be unsatisfactory, for resinous electricity could be produced on glass by rubbing it with fur, and, moreover, any object of the type we should now call an insulator could receive a charge when it was rubbed. Metals, it appeared, could not. So insulators were first called "electrics," and metals or conductors "non-electrics." This rather unlucky guess was given up when it was found that a metal could after all get a charge by friction. The charge could be held and detected successfully when the metal was fixed in an insulating handle. It was also found that when a charge of either kind was produced on an object A by rubbing it with an object B, then B invariably received a charge of the opposite kind. This fact inevitably suggested that one charge was due to the excess of some substance, and the other to the defect of this substance. It was arbitrarily supposed that the substance in excess produced a "vitreous" charge, and in defect produced a "resinous" charge. So a vitreous charge came to be known as "positive," and a resinous charge as "negative."

This guess was again unlucky, since we now suppose that a positively charged body is short of electrons,[1] and a negatively charged body has an excess of electrons. It was discovered also before the end of the seventeenth century that any set of substances could be arranged in order in a list, so that if any two on the list were rubbed together a positive charge would be produced on the higher one and a negative charge on the lower. Such a list is

Catskin,
Glass,
Silk,
Sulphur,
Flannel,
Ebonite,
Rubber,
Metals.

[1] See pp. 178–80 for discussion of electrons.

Induction of Charges

If the knob of the electroscope had been touched with a finger while the negatively charged body was near at hand, the leaves would have collapsed. If the finger had been removed and the negatively charged body kept where it was, the leaves might have stayed collapsed, or might have come up a little. If the negatively charged body had then been removed, the leaves would have stood up about as far as if they had been touched by the charged body.

The charge would then have been found to be positive; that is, opposite to the charge on the charged body.

This process is called "Charging by Induction."

The ordinary elementary explanation of this phenomenon is so inaccurate that it will not be given here, and the most satisfactory explanation will appear in Chapter IV, p. 50.[1]

If a charged body is used for charging by induction, its own charge is not in any way affected by the process. It may be used for giving hundreds of charges by induction, whereas if it had been giving charges by contact it would have been losing its own charge.

The reason that many charges may be obtained from one in this way is that mechanical work is done in separating opposite charges during the process, and the energy stored up by this work resides in the body charged by induction.

The "Electrophorus" is a good example of the use of charging by induction. It is simply a mechanical arrangement for producing a great number of charges quickly and fairly efficiently from a single original one. In the electrophorus, a piece of flat ebonite, called the *cake*, is fixed into a piece of metal, called the *sole*, and a flat metal disc, called the *plate*, is fixed to an insulating handle.

The cake is given a negative charge by beating it with warm fur, or rubbing it on your clothes, and is laid, facing upwards, on the bench. The plate is then brought in contact with the cake, earthed by the finger while in contact, and then removed. The plate should be found (unless the air is damp) to have a considerable positive charge, and a similar positive charge can be produced in this way many times from the same negative charge without appreciably reducing it.

The explanation of this is that, though the cake seems to

[1] Also see p. 43 for the function of the sole and p. 47 for the source of the energy.

be in contact with the plate all over, it is not really so, and though there is a negative charge on the plate in the few places where it makes contact, there is a much greater induced positive charge on its lower side at the many places where it does not make contact. When the plate is earthed the negative charge escapes and the positive charge is left.

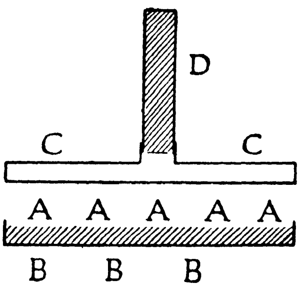


FIG. 3.

In Fig. 3, A is the cake, B the sole, C the plate, D the insulating handle.

Distribution of Charge on an Insulated Conductor

With the help of an electrophorus and an electroscope, simple experiments can easily be done to investigate the distribution of charge on a charged insulated conductor. A piece of pear-shaped wood, covered with tin-foil, fixed on an insulating stand, is a useful conductor to begin with. A small instrument called a "proof-plane" is also needed. It consists of a flat piece of metal of about 2 cm. diameter fixed to an insulating handle.

The conductor is given a charge by several applications of the electrophorus. The flat end of the conductor is touched by the plate of the proof-plane, which thus gets some of its charge. This charge is given to the electroscope by contact. The electroscope shows a deflection, which is observed. Electroscope and proof-plane are now discharged, and the proof-plane is brought in contact with the pointed end of the conductor. The charge is again given to the electroscope, and the deflection is found to be much greater. It is thus demonstrated that the proof-plane got more charge off the point of the conductor than off the flat part. The charge was taken off the point afterwards, so any effect of the gradual leakage of charge from the conductor should have produced the opposite result. The evidence thus indicates that there was at any moment more charge on the point than on the flat end. A better explanation of the behaviour of the proof-plane appears after capacity has been discussed.

Further experiments with miscellaneous conductors lead to the conclusion that the charge on a part of the surface varies with the convexity of the surface. The greatest density of charge is on sharp points. A more convex surface has more charge than a less convex one; any convex surface more than a flat one. A flat surface has more charge than a concave one, and a concave surface more charge than a more concave surface. Great density of charge tends to ionize, or electrify, the air molecules in the neighbourhood, and make the air conducting.

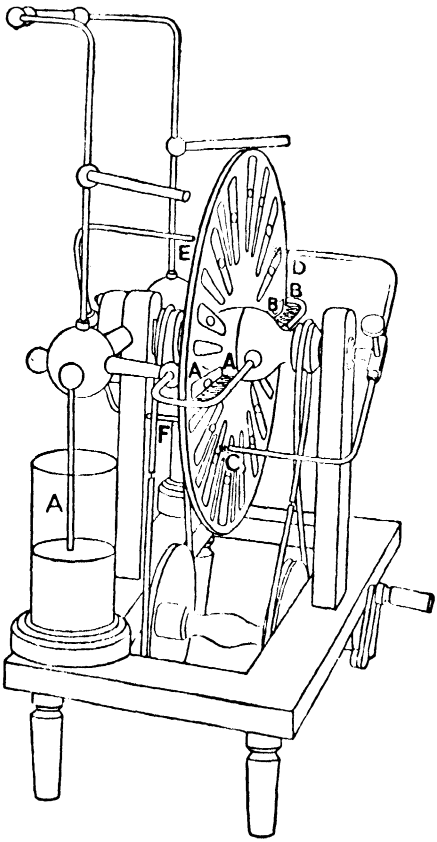


FIG. 4.

Charges thus tend to escape from points or to run on to points. For this reason electrical machines are provided with knobs rather than points; and for this reason lightning conductors are pointed.

No explanation of any of these phenomena can be given till (p. 43) after the discussion of "potential" and "capacity."

The Wimshurst Machine

The Wimshurst machine extends the method of the electrophorus, and employs the collecting power of points. It is

used to produce large electrostatic charges quickly. Its method of working is difficult to explain with words and a diagram only, but may be understood by following out the explanation below in detail with a real Wimshurst.

The Wimshurst machine consists of two glass discs, geared to revolve in opposite directions. On the outside of these discs, near the circumference, is a series of oval pieces of tin-foil, with hemispherical brass knobs on them.

There are four pairs of wire brushes, AA, BB, CD, EF (Fig. 4). AA and BB are connected respectively to the knobs of two Leyden jars A and B. The brushes CD are at opposite ends of a diameter-rod on one side of the plates, and the brushes E.F. at the opposite ends of a diameter-rod at the other end of the plates. AA and BB should perhaps be called "collecting combs" rather than brushes. They make no contact.

CD and EF make an angle of about 60° with each other and with the line AB.

Call the near-side disc, near CD, disc 1; and the far-side disc, near EF, disc 2. Let disc 1 be going round in an anti-clockwise direction, so that disc 2 is going round in a clockwise direction, both looked at from the near, or disc 1, side. Suppose now that one of the knobs on disc 1 between D and E happens to have a + charge on it. As it passes E, it will induce a − charge on the knob of disc 2 then touching E, and a + charge on the knob of disc 2 then touching F.

It is a good plan to mark these three charges +, −, and + respectively on the machine itself with different coloured chalks, for we are to follow their behaviour closely.

The original + charge is given up to the jar A via the combs AA. The − charge leaving E in disc 2 induces a + charge on a disc 1 knob at D, and a − charge on a disc 1 knob then at C; after which, it is given up to jar B via the combs BB.

The + charge leaving F also induces a − charge on a disc 1 knob at C and a + charge on a disc 1 knob at D, and is then given up to the jar A via the combs AA.

Now the − charge induced at C goes on its way, induces a + charge at F and a − charge at E, both on disc 2, and gives itself up to the jar B via BB.

This process continues indefinitely, the charges continually inducing fresh charges, and all + charges being given up to

jar A while all — charges are given up to jar B. A and B are thus charged up till their pressure is big enough to produce a spark. It should be noted that the charges on each disc increase cumulatively, since induction happens at both ends of CD and EF. The charge per disc reaches a maximum when the insulation in the neighbourhood begins to break down.

If a Wimshurst machine is worked in darkness, brush discharges can be seen from its discs when it has been turned several times; not immediately.

The whole process depends on having some accidental charge somewhere to start with. There usually is one; but if there is not, the machine can be started by holding a charged ebonite rod near the plates when they are revolving, thus inducing a starting charge.

The working of the Wimshurst machine depends, of course, on the ability of the combs AA and BB to collect charges from the charged discs passing them.

It has already been explained that the density of charge on a surface increases with the convex curvature of the surface. The points of the combs thus hold very dense charges which cause local ionization and enable the charges from the discs to leak on to the brushes.

The Wimshurst machine may be used for many attractive experiments. An ordinary small machine can be made to give a spark about 5 cm. long which requires a potential difference of between 50,000 and 100,000 volts between the knobs.

The brightness of the spark increases with the capacity of the Leyden jars which store the charge. The explanation of the action of a Leyden jar, and of the meaning of capacity, appears in Chapter III, p. 37. It is interesting to try the effect of turning a Wimshurst machine backwards. No spark is obtained. A careful following out of the sequence of events by the method just described will explain this.

The Electric Wind

If the Wimshurst is used to charge continuously an insulated conductor with a sharp point, a very distinct wind, called sometimes an Electric Wind, may be felt coming from this

point. The air molecules near the point are being ionized by the dense charge on the point, being given charges like that on the point. They are therefore repelled.

If a piece of wire in the form of a swastika (Fig. 5) with sharpened points is soldered to a central plate with a small dent at its centre of gravity, and is balanced in a horizontal plane on a vertical pointed rod fixed on an insulating stand, an "electric mill" is formed.

When the conductor is charged continuously by the Wimshurst, air molecules are repelled from all four points. The swastika is thus made to revolve away from its points. The one in Fig. 5 would thus revolve in a clockwise direction.

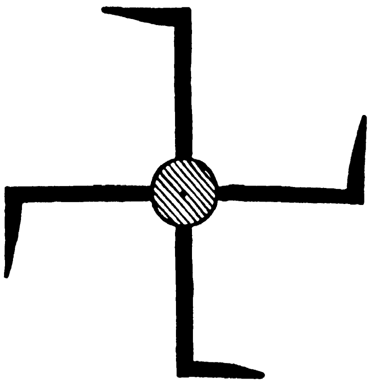


FIG. 5.

The hair of a person or a cat standing on an insulating platform may be made to stand up in a most satisfactory way with the help of a strong continuous charge from a Wimshurst, for individual hairs act as leaves of an electroscope. A catskin hung over one of the knobs shows this effect equally well. The repulsion of the electric wind ought to make the hair lie down, but apparently this effect is not strong enough to counteract the repulsion of individual hairs for each other.

Pith-balls covered with tin-foil (to make their surfaces conducting) may be made to perform very pleasing tricks near a Wimshurst. With careful manipulation they may be persuaded to experience repulsion the moment they have touched a charged Wimshurst knob which before attracted them; but the repulsion effect of like charges may be masked by the inductive effect, which causes attraction.

Faraday's Ice-pail Experiment

Before we begin the quantitative consideration of electric charges, we can learn much more about them from Faraday's

Ice-pail Experiment, which illustrates or implies almost all the important qualitative conclusions which can be reached about the behaviour of charged conductors.

This experiment gives us a dependable method of finding out which of two charged bodies holds the greater charge, and enables us to demonstrate:

(1) That the charge on a conductor containing no insulated charged bodies is wholly on its outer surface;

(2) That charges produced by friction are equal and opposite;

(3) That induced and inducing charges are equal and opposite, if the body on which the charge is induced completely surrounds the inducing charge;

(4) That if a hollow, closed insulated conductor contains insulated charged bodies, the charge on its outer surface is equal to the total charge inside it, and independent of the distribution of this charge.

This experiment was carried out by Michael Faraday with an old pail which was lying about in his laboratory. As the pail was ordinarily used for fetching ice, he called it his ice-pail, and so the experiment he used it for came to be known as the ice-pail experiment. It has nothing else to do with ice.

To carry out the experiment, one needs a large deep can, an electroscope, a metal ball with a piece of insulating silk thread attached to it, and a large slab of paraffin wax on which the can may stand.

The can is connected by a wire to the knob of the electroscope, and placed on the paraffin wax. The metal ball is given as large a charge as possible with a Wimshurst machine or an electrophorus.

There are three distinct parts to the experiment.

(1) The charged ball is lowered into the can. The induced charge on the can causes the leaf of the electroscope to rise gradually. When the ball is well inside the can (it is difficult to define the meaning of this exactly; but, if the diameter of the ball were half that of the can, it would be well inside when the top of it was a distance equal to its diameter below the mouth of the can) it will be found that moving the ball about inside the can has no effect at all on the divergence of the leaves.

(2) While the observer is watching the leaves closely, he allows the ball to touch the inside of the can for a moment. At the moment of contact the position of the leaves is unaltered. [They do, in fact, move a very little in some cases, if the can has an opening of appreciable size.]

Then the ball is removed (but not touched), the electroscope is discharged and disconnected, and the ball is brought up close to the electroscope to be tested for charge. It is completely discharged, and causes no vestige of a deflection on the electroscope.

(3) The electroscope is connected to the can again, and the ball is charged. The ball is made to touch the outside of the can. The electroscope is discharged and disconnected, and the ball is tested for charge as before. This time it is found to have a remaining charge.

Deductions from Faraday's Ice-pail Experiment.

Four distinct deductions may be made from this experiment.

(1) The distribution of the induced charge on the outside of a hollow conductor containing insulated charged bodies is independent of the position of such bodies, provided that they are completely inside the hollow conductor.

[This follows from the first part of the experiment.]

(2) The induced charge inside a hollow charged conductor completely enclosing insulated charged bodies is equal and opposite to the inducing charge.

[This follows from the fact that the touching of the ball against the inside of the can has no effect on the deflection of the electroscope. This is explained if the induced and inducing charges cancel out. A very small movement of the leaf at the moment of touching would be because the opening of the can was too big, and the body inside was not completely enclosed.]

(3) The total charge on the outside of a hollow conductor containing insulated charged bodies is equal to the total enclosed charge.

[Since the can was uncharged as a whole, the outside induced charge was equal and opposite to the inside induced charge, which was equal and opposite to the insulating charge.]

(4) THE CHARGE ON A CONDUCTOR IS ENTIRELY ON ITS OUTER SURFACE.

[The most important deduction. The charged ball gave up *all* its charge when, and only when, it touched the *inside* of the can.]

This result is of vital importance in the development of our knowledge of electricity. On this experimental fact the inverse square law of force between electric charges depends, though the experimental proof is due to Cavendish and Biot, rather than to Faraday.

It is possible to prove rigorously by pure mathematics that, if this conclusion is right, then the inverse square law must be true. This proof appears in the Appendix.

Faraday invented a very neat little experiment to demonstrate independently that the charge on a conductor is all on its outer surface. The apparatus is shown in Fig. 6.

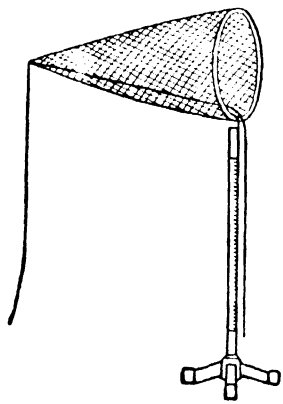


FIG. 6.

A conducting butterfly net on an insulating stand could be pulled inside-out by an insulating silk string.

A test with a proof-plane showed that whichever way out the net was the charge was all on the surface.

Faraday's ice-pail can also be used to show which of two charged bodies has the bigger charge. So long as both the charged bodies can get inside the pail, their size, shape and position inside the pail have no effect on the divergence of the leaves of an electroscope connected to the pail. The one with the bigger charge causes the bigger deflection.

We can now consider the law of force between charges, and the quantitative definition of charge. It was naturally supposed, by analogy with Newton's Law of Gravitation, that the Law of Force between electric charges was that of the Inverse Square. This was verified roughly but directly by Coulomb's Torsion-Balance Experiment, and accurately but indirectly by Cavendish's Experiment.

Coulomb's Experiment

In Coulomb's experiment two light balls are fixed at the ends of a light insulating rod, and this system is suspended by attaching a thin thread (preferably of quartz or silver) to the centre of mass of the system. One of the balls is given a charge, and a third insulated ball is charged and fixed in the same horizontal plane as the suspended balls, and near one of them. When the system is in equilibrium the couple due to the suspension is equal and opposite to that due to the electric force between the two charged balls, and hence this force can be calculated.

The apparatus is usually arranged so that the two balls which are to be charged just rest in contact, when uncharged, with no twist on the suspension. They are then given like charges, which are of approximately equal strength in order to minimize the error due to induction; and the calculation of the force operating is easy, since the couple due to the suspension for any twist may be taken as proportional to the angle of twist.

The errors in this experiment are larger than the divergence of its results from the Inverse Square Law, so that it may be regarded as direct evidence giving a certain amount of support to the law.

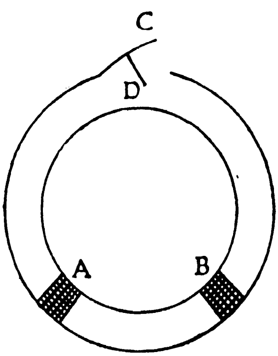


FIG. 7.

Cavendish's Experiment and the Inverse Square Law

Cavendish's experiment demonstrates that the charge on a conductor (containing no insulated charged bodies) is entirely on the surface. Its essentials are simple, but its results are conclusive.

Two conducting spherical shells are fixed together by an insulating ring AB (Fig. 7) so that they make no electrical contact anywhere. There is a trap-door C in the outer shell with a conducting arm D which makes contact with the inner shell when this is shut, and an insulating thread E (preferably

silk) attached to C by which C can be opened without discharging the outer shell at all. In the experiment the trap-door is shut, so that the inner and outer shells are in electrical contact, and the outer shell is given a charge. The trap-door is opened, so that the inner shell becomes insulated, and the outer shell is then discharged. The remaining charge on the inner shell is tested by connecting the latter through the trap-door to a sensitive electroscope; and this charge is found to be zero.

Since there has been no chance for the charge to escape, this test demonstrates that there was never at any time any charge on the inner sphere big enough to be detected by the electroscope.

The inverse square law may be inferred from this result with the help of either of two theorems, the mathematical proofs of which are given in the Appendix.

The first theorem, which follows from Gauss's Theorem, which is given in the Appendix, is that

"*It follows from the inverse square law that the charges on a conductor (containing no insulated charged bodies) are entirely on the outer surface.*"

The second is Cavendish's Proof of the Inverse Square Law, which shows that

"*If the charge on a spherical conductor (containing no insulated charged bodies) is entirely on the surface, then the law of force is that of the Inverse Square.*"

Maxwell was the final authority on Cavendish's experiment, and, from his estimation of the largest charge which his apparatus would have been unable to detect, he deduced that if the law were really that the force varied as $\frac{1}{r^{(2+\alpha)}}$ (where α was a small quantity), then

$$\alpha < \pm \tfrac{1}{21600}.$$

Quantity of Charge

Before we can get much further with the relation of charges to one another, we must find a way of thinking about quantity of charge numerically. The simplest way is this:

Imagine the stationary ball in the Coulomb Torsion Balance to be replaceable by another. Get two stationary balls which

produce separately equal deflections of the suspension. Such balls are defined to have equal charges. Keeping the distance between the stationary and moving balls the same by altering the zero of the suspension, observe the effect of putting the two stationary balls together. It will be found that within the limits of error of the experiment the force is doubled by doubling the stationary charge, and that it is quadrupled by doubling the moving charge as well. It is easy to see how, by such a series of experiments as this, Coulomb arrived at his fundamental law that if charges Q_1 and Q_2 are a distance r apart, the force between them is

$$A\,\frac{Q_1Q_2}{r^2},$$

where A is some constant, for if a charge Q_1 exerted at a given distance a certain force on a given charge, and another charge Q_2 exerted n times as much force on the same charge at the same distance, then Q_2 would be equal to nQ_1.

As it was found that the force between two charges depended also on the kind of medium that filled the space between them, Coulomb's formula for the force was modified so that it became

$$F = \frac{Q_1Q_2}{kr^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where k was a factor, depending on the nature of the medium. Our units will be chosen so that k is unity *in vacuo*, and for the present we shall consider everything as if it were *in vacuo*. We can now define exactly what we mean by some of the fundamental terms used to explain the phenomena of electrostatics. We will suppose that the meaning of charge is already clear enough to be used for other definitions.

Definition of Unit Charge

We can now define Unit Charge on the Electrostatic System. There are three electrical systems of units—the Electrostatic, Electromagnetic and Practical systems. A detailed account of these systems is given in Chapter IX, Part II. The unit of charge on the Practical system being the Coulomb, the unit on the Electrostatic is sometimes called the statcoulomb, but more frequently the Electrostatic or e.s. unit of charge.

An e.s. unit is defined as the charge which, when concentrated at a point, exerts a force of 1 *dyne on an equal point charge* 1 *cm. away* in vacuo. *It will appear later that* 1 *coulomb* $= 3 \times 10^9$ *e.s. units.*

It is vital to understand that all physical units must depend ultimately on the fundamental units of length, mass and time.

The Centimetre, the unit of length, is one-hundredth part of the length of the standard metre rod, kept at the Bureau des Poids et Mesures at Sèvres when the temperature of this rod is uniform at 0° C.

The Gramme, the unit of mass, is the mass of 1 c.c. of distilled water at a temperature of 4° C.

The Second, the unit of Time, is $\frac{1}{86400}$ of the length of a Mean Solar Day, which is the average length of a solar day taken throughout the year.

The unit of velocity is a velocity of one centimetre per second.

The unit of acceleration is an acceleration of one centimetre per second per second.

The definition of a unit of force is made possible by Newton's Second Law of Motion, which states:

"Rate of change of momentum is proportional to impressed force, and acts in the direction of that force."

Unit force, the dyne, is thus the force which produces unit rate of change of momentum.

Momentum is defined as the product of mass and velocity. So rate of change of momentum is the product of mass and acceleration, since acceleration is rate of change of velocity.

The dyne may therefore be more explicitly defined as the force which produces an acceleration of one centimetre per second per second in one gramme mass.

Simple Numerical Problems

Question 1.—What is the force between charges of 20 and 30 e.s. units 10 cm. apart *in vacuo*?

$$F = \frac{Q_1 Q_2}{r^2} \text{ dynes}$$

$$= \frac{20 \times 30}{100} \text{ dynes}$$

$$= 6 \text{ dynes.}$$

Question 2.—Charges of $+ 30$ and $- 10$ e.s. units respectively are on two similar pith-balls r cm. apart. The pith-balls are brought into contact so that they share their charge. They are then moved r cm. apart as before. What is the ratio of the force of repulsion between them now to that of attraction between them before they came into contact?

$$\text{Original force of attraction} = \frac{30 \times 10}{r^2}.$$

When they come into contact charges of $+ 30$ and $- 10$ combine to a single charge of $+ 20$. This is shared, giving each ball $+ 10$.

$$\therefore \quad \text{Force of repulsion} = \frac{10 \times 10}{r^2}.$$

$$\therefore \quad \frac{\text{Force of repulsion}}{\text{Original force of attraction}} = \frac{10 \times 10}{30 \times 10}$$
$$= \frac{1}{3}.$$

Experimental Work.

(1) With the help of charged ebonite and glass rods hung by thin suspensions, demonstrate that like charges repel and unlike attract. Use the charged rods to pick up light objects.

(2) Make or obtain a rough electroscope, using aluminium foil. Demonstrate the induction of charge, charging by contact, and charging by induction with it. Identify the sign of an unknown charge.

(3) Make or obtain an electrophorus, and use it to charge an insulated conductor.

Detect the charge with an electroscope.

Charge an insulated pear-shaped conductor with it, and use a proof-plane and electroscope to test the distribution of charge.

(4) Operate a Wimshurst machine. Try the effect on its spark of removing one or both of its Leyden jars. When the Leyden jars are replaced, use it to produce an electric wind and drive an electric mill. Make your own hair or someone else's stand up with it. Hang a catskin over one knob and charge it up.

Try the effect of bringing one or more pith-balls on long pieces of cotton near the Wimshurst when it is going.

Try the effect of working the machine backwards.

(5) Perform Faraday's ice-pail experiment. Use the ice-pail to compare the magnitude of charges, and demonstrate that frictional charges are equal and opposite.

This may be done by rubbing the end of an ebonite rod with a piece of fur on an insulating handle, putting each separately into the pail, and afterwards putting in both together.

CHAPTER II

INTENSITY, POTENTIAL, LINES OF FORCE, DISPLACEMENT

The Electric Field and Electric Intensity—Vectors and Scalars.—Potential—Definitions of Potential and Intensity—The Dielectric Constant—Potential due to a Single Point-charge—Potential Uniform on an Electrostatic Conductor—Earth Potential—Potential Reduced by the Approach of an Earthed Body—The Electroscope Measures Potential—Uncharged Conductor in an Electric Field—Lines of Force—Tubes of Force—Electric Displacement, or Induction.

The Electric Field and Electric Intensity

THE neighbourhood round an electric charge is called an Electric Field, and clearly other electric charges introduced into an electric field will have forces acting on them. Thus we may also say that when a force acts on an electric charge, that charge is in an electric field.

The intensity of the field at any point (usually known as the Electric Intensity E) is numerically equal to the force in dynes acting on 1 e.s. unit of charge (or statcoulomb) situated at the point. Its direction is the direction of the force on a positive charge. E, the electric intensity, is thus a quantity having both magnitude and direction, and thus it differs from Q, the charge, which has magnitude only.

Vectors and Scalars

All quantities in Physics are either vectors or scalars. Vectors are quantities like E which have magnitude and direction. [E is often called the "Electric Vector."] Other vectors in Physics are Velocity, Acceleration, Force, Momentum, and Length.

Scalars are quantities having magnitude only. Charge is thus a scalar. Other scalars are mass, volume, energy, and potential (which is to be explained next).

If two vectors are possessed by a body, or act on it together,

their magnitudes do not in general add up unless their directions happen to be the same. If two people try to push a large stone at the same time the total force on the stone will only be the sum of the forces exerted by each person if they are both pushing in the same direction. If they are pushing in different directions, the resultant force (which may be regarded as the vector-sum of the two original forces) is less in magnitude than the sum of the two magnitudes, and different in direction from either direction. Sums of vectors are thus obtained by the parallelogram law of elementary statics.

The magnitudes of scalars add up directly according to the ordinary rules of algebra.

The mathematical treatment of vectors, known as Vector Analysis, is a most powerful weapon in electrical theory. It is feared by beginners now, just as calculus was feared by beginners in Physics in the nineteenth century; but it will probably soon be regarded as necessary at an early stage, as calculus is regarded as necessary at an early stage now.

Potential

Let us consider two points A and B anywhere in space, separated by a distance of x centimetres. Suppose an electric intensity E (the same all the way) acts from A to B.

Let us now imagine that we have to bring a unit positive charge from B to A against the electric intensity. A force of E dynes will be pushing it back towards A all the time. So we shall have to exert a force of E dynes to get it along at all. In taking it the x cm. from B to A we shall thus do Ex ergs of work. This work will be the same whatever path we take. It does not matter whether we go straight from B to A or describe a big curve.

There are two ways of seeing that this is true.

By the direct method we see that since E acts directly from A to B it will require no work to move in any direction perpendicular to AB.

But any motion may be regarded as the sum of two component motions—one parallel to AB and the other perpendicular. The sum of all the parallel components can only be x whatever path we choose; and as these components alone require work the total work will be independent of the path.

The other method of argument is prettier.

Suppose that the work done can vary with the path. Let us go from B to A by one path doing work W_1. Let us then be pushed back to B by another path having work W_2 done on us. Everything will then be back exactly at the *status quo ante*, except that we shall have done work $(W_1 - W_2)$ and have nothing to show for it. Unless $(W_1 - W_2) = 0$, or rather $W_1 = W_2$, the principle of the conservation of energy will have failed. Therefore $W_1 = W_2$, and W is the same whatever path is chosen.

Now the quantity Ex, the work done in taking unit charge from B to A against the intensity E, depends only on the relative positions of B and A, and on the magnitude and direction (relative to AB) of the electric intensity E.

This quantity Ex is called the "Potential Difference" between B and A.

Let us now alter the conditions: Suppose there is a charge of Q units at A, and nothing at B; and suppose that all the other electric charges in the universe are so far away that the intensity due to them is negligible near A and B.

Now let us take a unit positive charge and go to infinity with it. This sounds unnecessarily enthusiastic. But we need only go far enough away for the charge at A to have no appreciable effect in our new neighbourhood. Let us then bring our charge back to B.

At first we shall do no work to move it. Presently the influence of the charge at A will begin to make itself felt. By the time we have got to B we shall have done a finite quantity of work. This quantity of work, the work done per unit charge against the electric intensity to bring a charge from infinity to B is known as the "Potential" of B.

Definitions of Potential and Intensity

We can now define Potential and the unit of Potential. The potential of a point is the work done per unit positive charge in bringing a charge from infinity to the point against the electric intensity.

"*The Electrostatic Unit of Potential (sometimes called the statvolt) is the potential of a point when* 1 *erg of work is done in*

bringing unit positive charge from infinity to the point against the electric intensity."

It will appear later that 1 statvolt = 300 volts.[1]

Before we give a final definition of electric intensity, let us consider more closely the relation between potential and intensity.

In our original case of the uniform intensity between A and B we saw that

Potential Difference between A and B = Intensity × distance AB.

If we turn this equation round we get

$$\text{Intensity} = \frac{\text{Potential Difference between A and B}}{\text{Distance AB}}$$
$$= \text{Drop of Potential per centimetre distance}$$
$$= \text{Negative Gradient of Potential.}$$

For the gradient of a hill is the change of vertical level per unit horizontal distance.

We can now give satisfactory definitions of intensity and unit intensity.

"*The electric intensity at a point is the force per unit charge at the point, and its direction is that of the force on a positive charge. It may also be regarded as the negative gradient of the potential in e.s. units per cm.*"

"*Unit electric intensity is the intensity when 1 e.s. unit of charge experiences a force of 1 dyne; and when the negative gradient of the potential is 1 e.s. unit per cm.*"

The Dielectric Constant

It was noticed in Coulomb's experiment that the force between two charges depended on the medium between them. The Dielectric Constant (sometimes called the Specific Inductive Capacity) of a medium is the ratio of the force between two charges *in vacuo* to the force between the same two charges the same distance apart in the medium. It is called k. Since a medium always diminishes the force k is always greater than unity for anything except a vacuum, for which it is unity.

[1] See Part II, Chapter IX, especially pp. 474–7.

Since the force F between two charges Q_1 and Q_2, r cm. apart, is given by

$$F = \frac{Q_1 Q_2}{kr^2}$$

and, by the definition of intensity, the force on Q_2 in the field E due to Q_1 is given by

$$F = EQ_2,$$

it follows that $$E = \frac{Q_1}{kr^2}.$$

Potential due to a Single Point-charge

Thus the intensity acting against us as we bring a unit positive point-charge from infinity to a distance of, say, x cm. from the charge Q, gradually increases from zero to $\frac{Q}{kx^2}$, but is neither constant nor increasing at a uniform rate. We cannot, therefore, calculate the potential at a distance x away by ordinary arithmetic. Though the problem is extremely simple (it is done in two lines (p. 316) at the beginning of Part II) it can only be done very clumsily without Integral Calculus.

For this chapter we must assume the result, which is a very simple one. The potential at a distance of x cm. is $+\frac{Q}{kx}$ e.s. units.

Numerical Questions about Potential.

Question 1.—Charges of 12, 36, 48 e.s. units are at the vertices of an equilateral triangle with sides of 6 cm. length. The dielectric constant is 6. Find the potential at the point in the centre of the triangle equidistant from all three charges.

The distance of the centre point is easily shown to be $2\sqrt{3}$ cms.

The potential of this point is the sum of the potentials due to the three charges taken separately. It is thus

$$\left(\frac{12}{6 \times 2\sqrt{3}} + \frac{36}{6 \times 2\sqrt{3}} + \frac{48}{6 \times 2\sqrt{3}}\right) \text{ e.s. units.}$$

or $$\frac{8}{\sqrt{3}} \text{ e.s. units.}$$

Question 2.—The potentials of two parallel plates, separated by a medium of dielectric constant 4, are 40 and 100 e.s. units respectively. The plates are 5 cm. apart. The intensity may be taken as constant in between them. Find (*a*) the value of the intensity, (*b*) the force on a charge of 2 e.s. units placed at any point between the plates.

(*a*) The intensity is the negative gradient of the potential.

$$\therefore \quad \text{Intensity} = -\frac{40 - 100}{5} \text{ e.s. units per cm.}$$
$$= 12 \text{ e.s. units per cm.}$$

(*b*)
$$\text{Force} = \text{Intensity} \times \text{charge}$$
$$= 12 \times 2 \text{ dynes}$$
$$= 24 \text{ dynes.}$$

Potential Uniform on an Electrostatic Conductor

An "electrostatic conductor" is a conductor in which no currents are flowing.

The fact that such a conductor is at the same potential all over follows from a very simple argument as follows:

(*a*) All the charges on the conductor are at rest.

(*b*) Therefore no intensity acts anywhere on the conductor.

(*c*) Therefore the potential has no gradient anywhere on the conductor.

(*d*) Therefore the potential is the same at all points on the conductor.

Earth Potential

It has been shown that the potential of a conductor whose charges are at rest is the same all over. The earth may be regarded as such a conductor, and it is so much bigger than any other conductors near it that its total charge, and consequently its total potential with respect to infinity, is not appreciably altered when such charges as we can deal with enter or leave it. Consequently we generally assume that "zero potential" means "earth potential" rather than "potential at infinity," and for all practical purposes the definition of potential should be altered to allow for this interpretation.

Actually the absolute potential could be of no interest to us unless another planet or star came close to us; and only differences of potential between neighbouring points really matter to us; so it does not matter for practical purposes where we take our zero, and we may as well for our own convenience take it at the earth.

Potential Reduced by the Approach of an Earthed Body

If an earthed body B approaches a positively charged body A, a negative charge is induced on B. If an exploring unit positive charge is now brought to A from infinity (or from earth) the negative charge on B will attract, and thus counteract to some extent the repulsion by A.

The work done in bringing the charge to A is thus reduced, so that the potential of A is reduced.

It should be noted here that an earthed body may be charged. B remains earthed, but has a negative charge.

We can perhaps see better how this odd arrangement occurs. When B is brought near A there is an electric intensity acting on it away from A. Before the steady state is reached, the positive charge on B is repelled from, and the negative charge attracted to, the surface near A.

Again, B, being near A, would be at positive potential if the charge on A were the only one present. So in order to succeed in remaining at zero potential B has to have a negative charge.

It is worth noting that the fact that positive charges move in the direction of the intensity may be usefully looked at in another way. Since the intensity is the negative gradient of the potential, it follows that positive charges always tend to move down the potential gradient; that is, from places of higher to places of lower potential. Negative charges, of course, tend to move in the opposite direction.

The Electroscope Measures Potential

If the leaves of an electroscope hold a positive charge (whether absolute or induced), they are at a positive potential to earth. They therefore move down the potential gradient —*i.e.* toward the case—until their weight pulling them downwards counteracts the electrostatic attraction pulling them outwards and upwards. If they are negatively charged, the

potential gradient from them runs up to earth, and they move up it toward the case.

If an earthed body, such as the hand, is brought near an electroscope, the potential of the leaves is reduced, though the charge on the electroscope is unchanged; so the divergence of the leaves is decreased. It is thus clear that variations of potential, rather than variations of charge, affect the position of the leaves. If the case is an earthed conductor, the potential gradient is steeper, and the electroscope consequently more sensitive, than if the case is simply an insulator like glass.

Moreover, insulators like glass tend to get stray charges on them which cause irregular effects.

Since it is only the potential gradient that matters, an electroscope should work equally well if we earth the knob, and insulate and charge the case. And so it does.

Uncharged Conductor in an Electric Field

The behaviour of the free charges on a conductor in an electric field, when its total charge is zero, follows the same lines.

Suppose AB is an uncharged conductor situated between a positive charge Q and an earthed plate E (Fig. 8). Let the line-graph show the potential as it was before AB appeared, falling continuously from V to zero in the distance x from Q to E.

Since the intensity is in the direction in which the potential is falling, it acts in the direction QE. When AB is placed in the field, the intensity drives the + charges to B and the − charges to A.

This process continues until the intensity due to the separation of charges on AB exactly counteracts the original intensity so that there is zero intensity along AB, which is thus at the same potential all over. So long as it is not at the same potential there will be an intensity in it which will continue to move the free charge.

The effect of the positive charge on B, numerically equal to the negative charge on A, and nearer to E than this negative charge, would be to raise the potential of E. A momentary intensity in E is thus produced, and this piles up a little more negative charge on its surface until its potential is again zero.

After the steady state has been reached, the fall of potential from Q to E is shown by the dotted graph. The conditions this graph must fulfil are that Q and E must be at the same potential as before (for it is assumed that Q is large enough to have its potential not appreciably lowered by the presence of AB), and that the potential along AB must be the same.

It follows from the graph that the intensity from Q to A and B to E must now be greater than it was before.

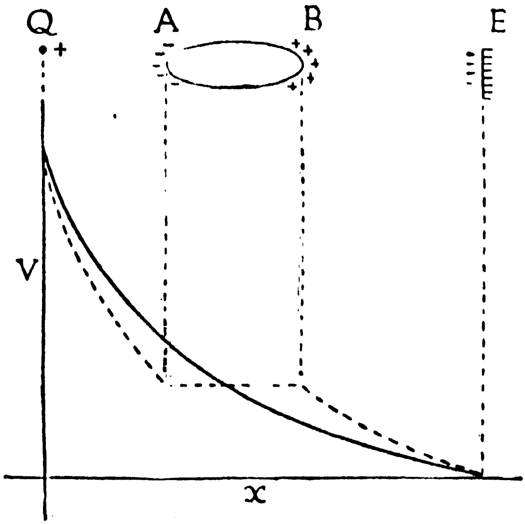


FIG. 8.

Lines of Force

Faraday postulated Lines of Force and Tubes of Force in order to be able to account for electrostatic effects without having to believe in "Action at a Distance."

For some problems they are very useful weapons of attack; particularly problems about electrostatic and electromagnetic fields. A Line of Force may be defined as follows:

1. "A Line of Force is a line tracing out such a path that the electric intensity is tangential to it at every point along it." Other properties follow from this definition.

2. Every line of force begins on a positive charge and ends on a negative charge. (This follows because the intensity acts from a positive charge to a negative charge.)

3. Lines of force cut equipotential surfaces normally. (An equipotential surface is a surface taken through points at the same potential. There is thus no potential gradient *along* such a surface. So the intensity must be perpendicular to an equipotential surface everywhere.)

4. Lines of force enter electrostatic conducting surfaces normally. (Conducting surfaces are equipotential surfaces, because conductors on which the charges are at rest are at the same potential all over.)

Two properties are qualitative, and may be inferred by looking at the distribution of lines of force.

5. Lines of force behave as if they were in tension along their length.

6. Lines of force repel each other in a direction perpendicular to their length.

One property is deliberately given to lines of force to make them useful in problems involving numbers.

7. The electric intensity at any point *in vacuo* is numerically equal to the number of lines of force per square centimetre of area, this area being taken so that the intensity at the point is normal to it.

Two more properties appear when we are dealing with lines of force in a material medium.

8. Lines of force are continuous across the boundary between a vacuum and a medium, or between two different media. (This follows from Property 2.)

9. In a medium of dielectric constant k the electric intensity due to given charges is $\frac{1}{k}$ of what it would be *in vacuo*. So the electric intensity is $\frac{1}{k}$ of the number of lines of force per sq. cm.

The distribution of lines of force between a point-charge and an earthed conducting plane are shown in Fig. 9. The effect of introducing an insulated uncharged conductor AB is shown in Fig. 10.

Electric lines of force can be plotted on glass, or on any dry insulator, with small particles of fibrous gypsum.

These particles are quite satisfactory if they are scraped off a lump of fibrous gypsum with a knife.

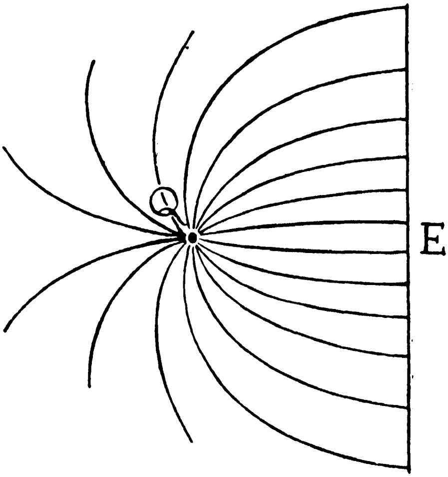


FIG. 9.

The figures obtained are quite recognizable, though not so good as the magnetic lines of force obtained with iron filings.

A field of several hundred volts per cm. is needed. The effects of Figs. 11 and 12, in the next chapter, may be obtained from 200-volt mains by using pieces of thin wire two or three millimetres apart on dry glass. Gentle tapping will usually show the formation.

Figs. 9 and 10 require a Wimshurst machine to show them.

One knob of the Wimshurst can be put against a perfectly horizontal dry glass plate to represent the point-charge, and the other connected to a piece of foil with a straight edge to represent the earthed plane.

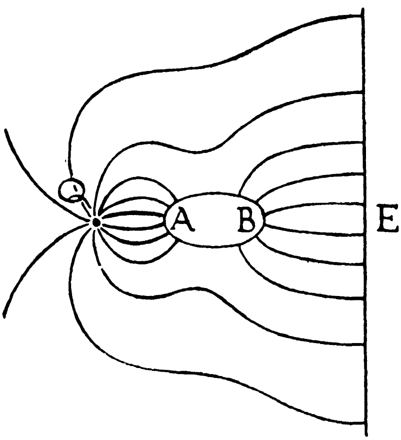


FIG. 10.

The figures can be shown with the knob several centimetres from the foil. If a wire ring is laid on the glass between the knob and the foil, the absence of lines of force inside a closed conductor can be shown.

The filings of gypsum inside the ring are quite unaffected by the external field.

Everyone should do, or at least see, these experiments if possible. They give a most valuable impression of the reality of the electric field, and of the behaviour of the lines of force.

Tubes of Force

A Tube of Force is the set of lines of force leaving unit charge.

In order to see how many lines of force there are in a tube of force, consider a unit positive point-charge *in vacuo.* Imagine a sphere of radius 1 cm. drawn round it. At every point on this sphere the intensity will be 1. (For intensity $= \frac{Q}{r^2}$, and Q and r are both 1.)

Therefore there will be 1 line of force per square cm. of surface of the sphere. But the area of the surface of a sphere of radius 1 cm. is 4π sq. cm.

So there are 4π lines of force in a Tube of Force.

Lines of force and tubes of force now give us directly the value of the electric intensity just outside the surface of a charged conductor *in vacuo* on which the charge per sq. cm. is σ.

Since there is no intensity inside the conductor, and all the lines of force must leave it normally, we must have σ tubes of force, or $4\pi\sigma$ lines of force, leaving each sq. cm. So the electric intensity just outside the surface is $4\pi\sigma$.

It may seem strange that both lines of force and tubes of force are used, whereas in Magnetism only lines are generally used. Moreover, D, the Electric Induction, which is explained in the next paragraph, is the number of TUBES of force per sq. cm., or $\frac{1}{4\pi}$ of the number of Lines per sq. cm., whereas B, the corresponding magnetic quantity, is the number of lines.

This inconsistency exists because electrical theory happened to develop as it did. Faraday was interested in electric intensity primarily, and required the tube of force because the unit charge can exist by itself. One tube of force is thus required as leaving one unit charge.

Maxwell, on the other hand, dealt with the electromagnetic

field as a whole. The isolated magnetic pole does not exist, and there is thus no need for the conception of a tube of magnetic force leaving a single pole.

There is thus a rather unfortunate lack of symmetry in the mathematical treatment of electric and magnetic field and induction. It may be long before the trouble is cleared up.

Electric Displacement, or Induction

This quantity, D, is defined as follows:

The Electric Displacement, or Induction, D, is the number of tubes of force per sq. cm. taken normal to the surface. Since there are 4π lines per tube, and since $E = \frac{1}{k}$ of number of lines per sq. cm., it follows that

$$D = \frac{kE}{4\pi}$$

and

$$E = \frac{4\pi D}{k}.$$

This quantity has been of the greatest importance in the development of Physics, for Maxwell developed the Electromagnetic Theory of Light by considering its rate of change, and it may thus be said to have led to the discovery of wireless waves. It is also very useful in certain other problems. It is sometimes alternatively called the "Electric Strain." The energy per unit volume in a dielectric can be shown to be $\frac{1}{2}ED$, which is obviously analogous to the energy per unit volume of an elastic solid under the action of a stretching force or stress, in which the energy per unit volume is $\frac{1}{2}$ stress $\times$ strain. Strain is in this case the distortion, or change of volume per unit volume, and stress is the force per unit area.

CHAPTER III

CAPACITY, ENERGY AND MECHANICAL FORCE

Capacity—Capacity of a Charged Sphere at Infinity—The Parallel-plate Condenser—Types of Parallel-plate Condensers—Condensers in Series and Parallel—Energy of a Charged Condenser—Distribution of Potential and Charge—Brush-discharge and Discharge from Points—The Electrophorus—Mechanical Force on the Surface of a Charged Conductor—Attraction of Uncharged Bodies by Charged Bodies—Attraction between the Plates of a Condenser—Miscellaneous Electrostatic Problems.

Capacity

THE capacity of a conductor is defined as the ratio of its charge to its potential: $C = \frac{Q}{V}$. Unit capacity is thus the capacity of a conductor which holds a charge of 1 statcoulomb when its potential is 1 statvolt. It may be called a statfarad. We have seen that the potential of a conductor depends not only on its own charge, but also on the distribution of all charges, conductors and dielectrics round it. It follows, therefore, that the capacity of the conductor depends not only on the size and shape of the conductor itself, but also on the distribution of all charges, conductors and dielectrics round it.

Capacity of a Charged Sphere at Infinity

In general the calculation of the capacity of a conductor is a difficult matter, involving much mathematics. But one capacity-problem is easy to solve.

Let us consider the arrangement of the lines of force round a uniform charged sphere so far from other charged bodies that their fields are negligible. [This is the meaning of the phrase "a charged sphere at infinity."]

By symmetry, the lines of force will go out uniformly in all directions in straight lines radiating from the sphere's centre.

Since all the charge of the sphere is at the surface, the lines begin at the surface, leaving it normally.

The arrangement of lines of force outside the surface is thus in every respect the same as if they had come from the whole charge on the sphere concentrated at the centre. Let this charge be Q, and the radius be r.

Since this is true of the lines of force, it must also be true of the Intensity at all points.

Since the potential of the sphere is the same at all points, we need only bring our unit charge from infinity to the surface

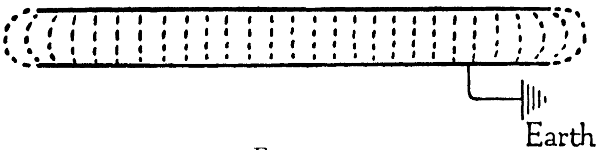


FIG. 11.

to find the potential. The potential at the surface is thus exactly the same as the potential at a distance r cm. from an

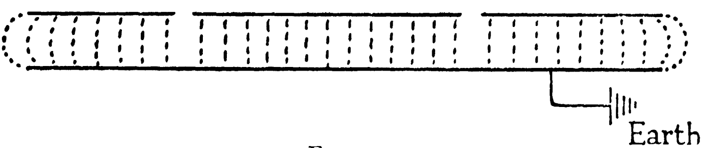


FIG. 12.

isolated point charge Q. As we saw in the last chapter this potential, *in vacuo*, is $\frac{Q}{r}$.

$$\text{So the capacity} = \frac{\text{charge}}{\text{potential}}$$
$$= \frac{Q}{Q/r}$$
$$= r.$$

The capacity of a sphere at infinity *in vacuo* is thus numerically equal to its radius.

The Parallel-plate Condenser

When a conductor has its capacity increased by the introduction of another conductor, usually earthed, to its neighbourhood, the arrangement is known as a Condenser. The most usual kind of condenser consists of two parallel plates separated by a medium of high dielectric constant.

The conception of lines of force makes the calculation of the capacity of such a condenser a simple matter.

The lines of force usually take the form shown in Fig. 11. Since the lines of force do not go straight across at the edges the intensity is irregular there, and the capacity is consequently too difficult to calculate. Where very accurate capacities are needed the condenser is arranged as in Fig. 12, the insulated plate being surrounded by another insulated plate, called a guard-ring. The central plate alone takes the measured charge, though inner and guard-ring plates are kept at the same potential.

In calculating the capacity we shall assume that the lines of force go straight across uniformly at every point.

Let our condenser have area A sq. cms., its plates being d cm. apart, separated by a medium of dielectric constant k. Let the charge on the insulated plate be Q, and the potential of the insulated plate V.

There are then $\frac{Q}{A}$ units of charge on each sq. cm. of the insulated plate. A tube of force, consisting of 4π lines of force, leaves each unit charge. There are thus $\frac{4\pi Q}{A}$ lines of force per square cm. between the plates. The electric intensity is therefore $\frac{4\pi Q}{kA}$.

The work done in taking unit charge along the distance of d cm. between the plates against this Intensity is therefore $\frac{4\pi Qd}{kA}$ ergs.

The potential V is therefore given by

$$V = \frac{4\pi Qd}{kA}.$$

The capacity C is then given by

$$C = \frac{Q}{\frac{4\pi Q d}{kA}}$$

$$= \frac{kA}{4\pi d}.$$

Numerical Question.

Question 1.—A "wireless" variable condenser consists of 11 plates, each of area 15 sq. cm. uniformly spaced. The odd plates are insulated, the even plates earthed. The dielectric

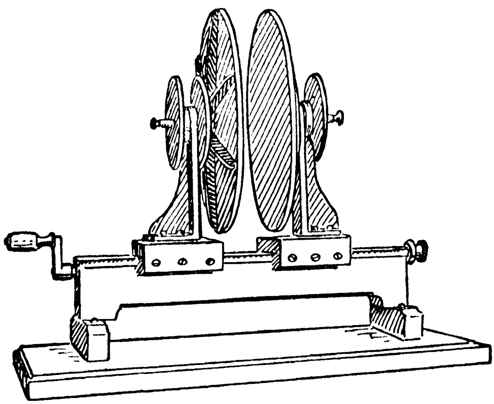


FIG. 13.

is oil of dielectric constant 4. The distance between successive plates is 1 millimetre. Find the maximum capacity.

There are 10 gaps between the plates. The capacity is thus the same as if there had been two plates one millimetre apart of area 150 sq. cm. The capacity is thus given by

$$C = \frac{4 \times 150}{4\pi \times 0 \cdot 1} = \frac{1500}{\pi} \text{ e.s. units.}$$

Now the farad is the capacity of a body which holds a charge of 1 coulomb when its potential is 1 volt. 1 coulomb

is 3×10^9 e.s. units; 1 volt is $\frac{1}{300}$ e.s. units. So 1 farad is $\frac{3 \times 10^9}{1/300}$, or 9×10^{11} e.s. units.[1]

[It is thus the capacity of sphere of radius 9×10^{11} cm. The radius of the earth is $6{\cdot}4 \times 10^8$ cm. Hence the capacity of the earth is about 0·0007 farad. So in this case the "practical" unit is not very practical.] The microfarad is thus 9×10^5 e.s. units. Hence the capacity of our condenser would be $\frac{1500}{9 \times 10^5 \pi}$, or about ·0005 microfarad.

Types of Parallel-plate Condensers

The simplest kind of condenser consists of two parallel plates, one fixed and one movable, so that the distance between the plates can be accurately adjusted (see Fig. 13).

The more usual variable condenser is the type used in wireless sets, which has two groups of parallel plates, each group of plates being connected to each other. The plates are arranged alternately. The area of plate common to both can be varied by turning one group of plates about a central spindle.

The fixed condenser of large capacity (or small capacity) in a wireless set consists of alternate plates of tin-foil and very thin mica. The large capacity is obtained partly by making A large (many plates), but chiefly by making d small (very thin mica plates).

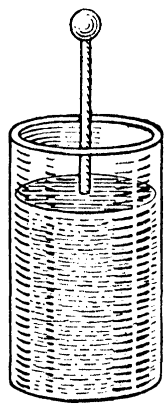


FIG. 14.

The famous Leyden jar (Fig. 14) consists simply of a glass jar (the thinner the glass the better) coated with tin-foil outside and inside. Contact is usually made inside by a small chain attached to the discharge knob. The outside coating is earthed by standing the jar on the table.

The advantage of the Leyden jar is not that its capacity is very large (it is not), but that it will stand very large P.D.'s, of the order of 50,000 volts, without breaking down.

Numerical Questions.

Question 2.—Find the capacity, in microfarads, of a cylindrical Leyden jar 20 cm. high and 10 cm. in diameter, the glass being 1 mm. thick and of dielectric constant 6.

$$\text{Area} = 2\pi rh + \pi r^2 = \pi r(2h + r) = \pi \times 5 \times 45 \text{ sq. cm.}$$

$$\text{Capacity} = \frac{6 \times (\pi \times 5 \times 45)}{4\pi \times 0{\cdot}1}$$

$$= 3375 \text{ units}$$

$$= \frac{3375}{9 \times 10^5} = 0{\cdot}00375 \text{ microfarad.}$$

Question 3.—A mica fixed condenser, of 10 sq. cm. surface area and 1 cm. thickness, is made of alternate sheets of mica and foil, 0·01 cm. thick each. The dielectric constant of mica is 7. Find the capacity in microfarads.

The number of mica sheets is $\frac{1{\cdot}00}{0{\cdot}02}$, or 50. So the effective area is 500 sq. cm.

$$\text{Thus the capacity} = \frac{500 \times 7}{4\pi \times 0{\cdot}01} \times \frac{1}{9 \times 10^5} \text{ microfarads}$$

$$= 0{\cdot}031 \text{ microfarad.}$$

Condensers in Series and Parallel

It is easily shown that if C be the capacity of a number of condensers C_1, C_2, C_3 . . . etc., in parallel, then

$$C = C_1 + C_2 + C_3 + \ldots$$

and that if C' is their capacity in series, then

$$\frac{1}{C'} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots$$

For consider them first in parallel (Fig. 15), charged to a common potential difference, V, between the plates; then the charges on the insulated plates are VC_1, VC_2, VC_3, VC_4, etc. The total charge on the insulated plates is then

$$V(C_1 + C_2 + C_3 + C_4 + \ldots)$$

Hence the capacity of the condensers in parallel

$$C = \frac{\text{charge}}{\text{potential}} = \frac{V(C_1 + C_2 + C_3 + C_4 + \ldots)}{V}$$

$$= C_1 + C_2 + C_3 + \ldots$$

Consider them now as arranged in series (Fig. 16) and suppose a charge Q to be put on the insulated plate of C_1. Since all the lines of force leaving this plate arrive on the opposite plate, a charge of $-Q$ will be induced on the opposite plate; the first plate of C_2 will thus have a charge of $+Q$, since the whole conductor, consisting of the second plate of C_1 and the first of C_2, has no charge. The other charges will then be arranged as in the figure.

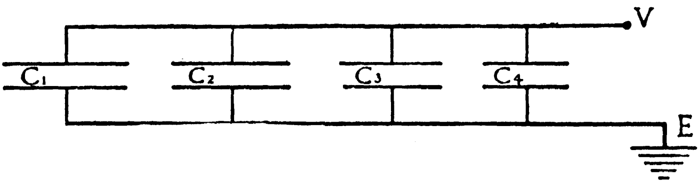


FIG. 15.

Hence if V_1, V_2, V_3 are the potential differences across the plates of C_1, C_2, C_3, etc., the total potential difference

$$V_1 + V_2 + V_3 \ldots = \frac{Q}{C_1} + \frac{Q}{C_2} + \frac{Q}{C_3} + \ldots$$

But the whole set taken in series may be regarded as one condenser carrying a net charge Q, so that the potential

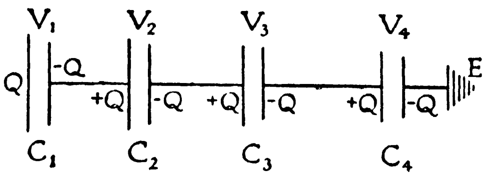


FIG. 16.

difference between the first plate of the first condenser and the last plate of the last will be $\frac{Q}{C'}$.

$$\text{Hence} \quad \frac{Q}{C'} = \frac{Q}{C_1} + \frac{Q}{C_2} + \frac{Q}{C_3} + \ldots$$

$$\therefore \quad \frac{1}{C'} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots$$

Numerical Question.

Question 4.—Condensers of capacity 4, 6, and 10 units are connected together (*a*) in parallel, (*b*) in series.

Find their combined capacity in each case.

In parallel their capacity $= 4 + 6 + 10 = 20$ units.

In series
$$\frac{1}{C'} = \frac{1}{4} + \frac{1}{6} + \frac{1}{10}$$
$$= \frac{15 + 10 + 6}{60} = \frac{31}{60}.$$
$$\therefore \; C' = \frac{60}{31} = 1\frac{29}{31} \text{ units.}$$

Energy of a Charged Condenser[1]

Work is done to charge a condenser, so it must contain energy. Let us imagine the case of a condenser charged to a potential V, when it holds a charge Q. However the charge was put there, the energy at the end is the same, and is equal to the work done to get the charge there. Suppose the charge is put on bit by bit. The first little bit of charge experiences no opposition, for there is no electric field to oppose it, but the last bit of charge experiences the opposition due to a potential V. So the work done on the first and last bits together is the same that would have been done if they had both been brought up against a potential of $\frac{V}{2}$.

The same is true of the second and last-but-one bits, and of the third and last-but-two bits. So the total work done is the same that would have been done to bring up a charge Q if the potential had stayed at $\frac{V}{2}$ all the time.

Thus the energy of the charged condenser is given by

$$E = \tfrac{1}{2}QV.$$

This argument only holds when V is proportional to Q.

Also, since C, the capacity $= \frac{Q}{V}$

and
$$Q = CV,$$
$$V = \frac{Q}{C}.$$

[1] See p. 321 for proof using calculus.

Substituting for Q in the energy-equation we get

$$E = \tfrac{1}{2}V^2C.$$

Substituting for V we get

$$E = \frac{\tfrac{1}{2}Q^2}{C}.$$

This energy is all transformed into heat, if the two plates are connected together, or if a spark passes between them which completely discharges the condenser.

Numerical Questions.

Question 5.—A condenser of capacity 4·2 e.s. units is charged to a potential of 1000 e.s. units. Find how many calories of heat are given up in a spark which completely discharges it.

$$\begin{aligned}
\text{Energy lost} &= \tfrac{1}{2}V^2C \\
&= \tfrac{1}{2} \times 4{\cdot}2 \times (1000)^2 \\
&= 2{\cdot}1 \times 10^6 \text{ ergs} \\
&= \frac{2{\cdot}1 \times 10^6}{4{\cdot}2 \times 10^7} \text{ calories} \\
&= 0{\cdot}05 \text{ calorie.}
\end{aligned}$$

Question 6.—A Leyden jar of capacity 0·02 microfarad is charged to a potential 30,000 volts. How many calories of heat can it give up on discharge?

$$\begin{aligned}
0{\cdot}02 \text{ microfarad} &= 0{\cdot}02 \times 9 \times 10^5 \text{ e.s. units.} \\
&= 1{\cdot}8 \times 10^4 \text{ e.s. units.} \\
30{,}000 \text{ volts} &= \frac{30{,}000}{300} = 100 \text{ e.s. units.} \\
\text{Ergs of energy given up} &= \tfrac{1}{2}V^2C \\
&= \tfrac{1}{2} \times 1{\cdot}8 \times 10^4 \times (100)^2 \\
&= 9 \times 10^7. \\
\text{Calories given up} &= \frac{9 \times 10^7}{4{\cdot}2 \times 10^7} = 2{\cdot}14 \text{ calories.}
\end{aligned}$$

[When a spark will pass between large brass knobs about 1 cm. apart, the P.D. is about 20,000 volts. So the amount of heat the average Leyden jar gives up when it sparks across 1 cm. is something like 1 calorie.]

Question 7.—Condensers of capacity 100 and 200 e.s. units are charged to potentials of 400 and 300 e.s. units respectively. They are joined in parallel.

What is their potential then, and what fraction of their energy is lost?

The central facts are that:

(1) No charge is lost when they are connected in parallel.

(2) Since their potentials are the same when they are in parallel, their charge must be shared in proportion to their capacity.

Their total charge ($Q = CV$) is thus

$$(100 \times 400 + 200 \times 300) = 100{,}000 \text{ units.}$$

Their combined capacity is

$$100 + 200, \text{ or } 300 \text{ e.s. units.}$$

So their potential $\left(V = \frac{Q}{C}\right)$ is

$$\frac{100{,}000}{300}, \text{ or } 333\tfrac{1}{3} \text{ e.s. units.}$$

$$\begin{aligned}
\text{Their initial energy} &= \tfrac{1}{2}C_1V_1^2 + \tfrac{1}{2}C_2V_2^2 \\
&= \tfrac{1}{2} \times 100 \times (400)^2 + \tfrac{1}{2} \times 200 \times (300)^2 \\
&= 8{,}000{,}000 + 9{,}000{,}000 \text{ ergs} \\
&= 17 \times 10^6 \text{ ergs.} \\
\text{Their final energy} &= \tfrac{1}{2}QV \\
&= \tfrac{1}{2} \times 100{,}000 \times \left(\frac{1000}{3}\right) \\
&= 16\tfrac{2}{3} \times 10^6 \text{ ergs.}
\end{aligned}$$

Thus $\frac{\frac{1}{3}}{17}$, or $\frac{1}{51}$ of the original energy is lost, or dissipated irrecoverably as heat.

Distribution of Potential and Charge

We can now see in a general way why the charge is distributed as it is on a conductor, and how a proof-plane works.

The potential all over the conductor is the same.

You must, therefore, need as much work to bring a point charge to the point A (Fig. 17) as to bring it to the flat surface BC. But at A most of the rest of the surface is much further away on the average than it is at BC. So the repulsion at A

must be much more due to the charge in the immediate locality. So this charge at A must be more concentrated. Thus though the potential all over an insulated pear-shaped conductor is the same, the charge per unit area is greater the more convex the surface, and is greatest at points. Though the proof-plane goes to the potential of the surface wherever it touches, its capacity can be shown to increase with the convexity of the surface, and so it picks up more charge at the same potential.

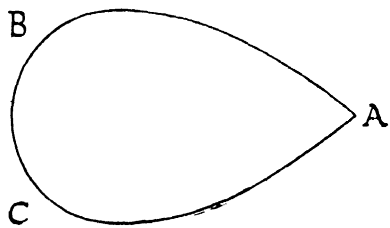


FIG. 17.

Brush-discharge and Discharge from Points

If the surface of a conductor is at a very high potential, or is very pointed, it has a great concentration of charge, and consequently a very strong electric field just outside. This field may pull the air-molecules apart, and so ionise them, or alternatively it may give them a charge like its own. Whichever the cause, a " brush-discharge," sometimes visible when there is not too much light, is one of the results, and an " electric wind," caused by the repulsion of the like-charged particles from the surface, is the other. The brush-discharge may be seen from high-power transmission lines at night; and points on a conductor are used in preventing it from reaching too high a potential, and in inducing lightning flashes to go to earth harmlessly through the pointed lightning-conductor. The point-discharge is best illustrated in the electric mill described in Chapter I.

The Electrophorus

The function of the " sole " in the electrophorus, which was not properly explained in the first chapter, since it involved potential, may be given now.

The sole forms an efficient earthed plate which, with the charged surface of the cake, forms a condenser of large capacity. The larger the capacity, the less the potential of the cake for a given charge. The less the potential the less chance there is of leakage to earth, and the larger the capacity, the larger the charge for a given potential.

The purpose of the sole is thus to prevent or minimise leakage of the charge to earth; or to obtain the largest possible charge for a given potential.

These statements express the same fact, but regard it from different points of view.

Mechanical Force on the Surface of a Charged Conductor

Let us suppose that our conductor has, at the point we are considering, a charge of σ units per sq. cm. σ is usually called the "surface density of charge."

However irregular the surface is, we may always consider an area so small that it may be regarded as flat and uniformly charged. Let such an area be α. Then the charge on it is $\sigma\alpha$.

The argument which follows, though short and simple, is for most people difficult to accept. The steps in it will therefore be numbered to make them easy to understand clearly. The difficulty should come in accepting, rather than understanding, the reasoning.

1. Since 4π lines of force leave the surface normally for each unit charge, $4\pi\sigma\alpha$ lines leave our area-element.

2. Though these lines represent the whole of the intensity, it cannot be true that the whole of the intensity is due to the local charge $\sigma\alpha$. Some of it must be due to the combined effect of all the other charges.

3. Force can only be exerted on the area $\sigma\alpha$ by the part of the intensity which is due to all the other charges. [The intensity due to a charge does not act on the charge itself.]

4. We must therefore determine how many of our $4\pi\sigma\alpha$ lines are due to all the other charges.

5. If the charge $\sigma\alpha$ were all by itself, lines of force would go out normally from *both* sides of its flat surface; $2\pi\sigma\alpha$ going outwards and $2\pi\sigma\alpha$ going inwards.

6. But since there is no intensity inside an electrostatic

conductor, the inward intensity must be exactly counteracted by an equal outward intensity from somewhere.

7. This can only come from all the other charges.

8. Therefore, of the $4\pi\sigma\alpha$ lines of force leaving the surface, $2\pi\sigma\alpha$ are due to the charge $\sigma\alpha$, and $2\pi\sigma\alpha$ due to all the other charges.

9. Thus $2\pi\sigma$ lines of force per square cm. are due to the other charges. So if the dielectric constant is k, the intensity acting on the charge σ per unit area is $\frac{2\pi\sigma}{k}$.

10. Therefore the mechanical force acting outwards on the charge is $\frac{2\pi\sigma^2}{k}$ dynes per sq. cm.

Attraction of Uncharged Bodies by Charged Bodies

The attraction of uncharged bodies for charged bodies can now be explained. There are two satisfactory explanations; one in terms of mechanical force, the other of energy.

The induced charges on an uncharged body are equal and opposite, since the body *is* uncharged. But since the electric intensity is greater on the side near the charged body, and is also proportional to the surface density of charge (for it is $\frac{4\pi\sigma}{k}$, as we have seen in the discussion of lines of force), it follows the charge on the near side is more concentrated.

From this it can be proved (*though on y very clumsily without calculus*) that the total mechanical force acting on the near side toward the charged body is greater than that acting on the far side away from the charged body. So the uncharged body is as a whole attracted to the charged body

The proof by considering energy is as follows. It is a corollary of the universal Principle of the Conservation of Energy that any system will tend to assume the configuration in which its potential energy is least.

We may regard our charged and uncharged bodies as forming two plates of a condenser, whose energy is $\frac{1}{2}QV$ if Q is the charge and V the potential difference between the two. Any movement of the two toward each other will leave Q unchanged, but will diminish V. It will thus diminish the

energy. By our principle this movement will therefore happen if it can.

Attraction between the Plates of a Condenser

The argument of the last section applies to the plates of a condenser. Suppose a parallel plate condenser of circular section and radius 10 cm. has its plates 1 millimetre apart, separated by oil of dielectric constant 6. If its potential difference is 3000 volts, or 10 stat volts, let us find the force of attraction between the plates.

Q, the total charge is the product of the potential and the capacity.

$$\therefore Q = V \times \frac{kA}{4\pi d}$$

$$= 10 \times \frac{6 \times \pi \times 10^2}{4\pi \times 0{\cdot}1}$$

$$= 15{,}000 \text{ e.s. units.}$$

$$\therefore \sigma, \text{ the charge per unit area} = \frac{15{,}000}{\pi \times 10^2} = \frac{150}{\pi}.$$

$$\text{Total force} = \frac{2\pi\sigma^2 A}{k}$$

$$= 2\pi \, . \, \frac{150 \times 150}{6\pi^2} \times \pi \times 10^2 \text{ dynes}$$

$$= 750{,}000 \text{ dynes.}$$

Miscellaneous Electrostatic Problems

Problem 1.—The plates of a parallel plate condenser are arranged vertically. One is insulated, charged, and connected to an electroscope, the other earthed. Explain what is observed when (*a*) the earthed plate is raised vertically; (*b*) the earthed plate is slowly moved away horizontally; (*c*) a sheet of ebonite is inserted between the plates; (*d*) the earthed connection of the earthed plate is broken, leaving the plate insulated.

(*a*) When the earthed plate is raised vertically, the effective area of the condenser is reduced, though the charge is unchanged. The capacity is thus reduced. So the potential increases. So the leaf of the electroscope rises further.

(*b*) When the earthed plate is moved away horizontally,

the capacity, $\frac{kA}{4\pi d}$, is reduced, since d is increased. So again the potential increases, and the electroscope leaf rises further.

(*c*) If a sheet of ebonite is inserted, k in the formula is increased. So the potential is decreased, and the deflection of the leaves becomes less.

(*d*) If the earth connection is broken, the like charge induced on the far side of the uncharged plate, together with any negative charge picked up by conduction, increases the potential of the insulated plate (thus decreasing the capacity). So the electroscope leaf rises further.

Problem 2.—Explain the electrical actions involved in the use of the electrophorus. What would be the advantage of increasing its area? Where does the electrical energy come from?

The cake is first charged by friction. The sole, being in contact with the table, is earthed, and forms a second plate to a condenser holding the charge. The potential of the charge is thus reduced, and the tendency to leakage of charge diminished. When the plate is in contact with the cake, it picks up some negative charge from the cake at its points of contact. But the negative charges in the hollows of the cake induce a much greater positive charge on the underside of the plate, and an equal negative charge on the top.

When the plate is earthed, all the negative charge on the plate escapes, but the positive induced charge remains. When the plate is again insulated it still holds all the positive charge, and takes this with it when it is removed. This operation has not in any way affected the original negative charge in the hollows of the plate, so that it is theoretically possible to repeat the operation an infinite number of times.

Actually the charge slowly leaks away, but of course this leak has nothing to do with the use of the electrophorus, but would act in any case.

Increasing the area of the electrophorus would proportionately increase the charge obtained in each operation.

The plate and cake form two plates of a condenser oppositely charged. There is therefore a mechanical force of attraction between them.

The plate is removed against this attraction; and work is done. This work separates the charges, and thus is stored as electrical energy.

Problem 3.—What happens to the potential of the plate of the electrophorus throughout the charging operation?

The potential, originally zero, is increased (in the negative direction) as the plate first approaches the cake, because the cake, being made of ebonite, gets a negative charge when rubbed.

When the plate is earthed, the potential becomes zero.

This is managed by the inflow, from the unlimited reservoir of charge formed by the earth, of a large positive charge, nearly the same as the total negative charge on the sole, to counteract the negative potential-forming power of this negative charge.

When the earthing connection is removed, the plate obtains a very small negative potential as the charge leaks from the cake. The plate now has a large positive charge but a small negative potential. As the plate is removed the induced effect of the cake diminishes, and the potential becomes large and positive. If the plate now touches the outside of an insulated conductor of large capacity it shares its charge, keeping only a quantity proportional to its own capacity. Its potential thus becomes small and positive. If it had touched the *inside* of an insulated conductor it would have given up all its charge, and fallen to the same small potential.

When it is finally removed its potential will be small and positive if it has touched the outside of the conductor, and zero if it has touched the inside of the conductor.

CHAPTER IV

ELECTROSCOPES AND ELECTROMETERS

Divergence of Leaves of Electroscope by Induced Charge—Divergence of Leaves of Charged Electroscope—Charging an Electroscope by Induction—The Condensing Electroscope—The Quadrant Electrometer—Adjustment of Quadrant Electrometer—Experiments with the Quadrant Electrometer: Sensitivity to Potential; Variation of Sensitivity with Needle Potential; Measurement of Alternating Potentials. Idiostatic Use of Electrometer; Comparison of Capacities—The Attracted-disc Electrometer.

Divergence of Leaves of Electroscope by Induced Charge

THE central rod and leaf of an electroscope, when a charged body is brought near its knob, may be regarded as a conductor in an electric field (see Fig. 18).

Suppose a positively charged body C is brought near the knob A. At the first instant, before the equilibrium position has been reached, there will be a potential gradient from A to B, so that the electric intensity will drive the positive charges from A to B, and the negative charges from B to A. This movement will cease as soon as the potential is the same all over AB. The leaves at B will now be positively charged, and in a potential gradient from B to the earthed shield of the electroscope. We may now consider:

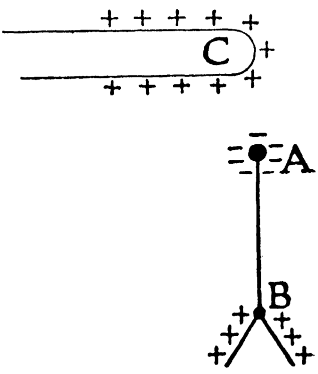


FIG. 18.

Either (1) That the leaves at B, being similarly charged, will repel each other;

or (2) That the leaves at B, being positively charged, will move down the potential gradient toward the earthed casing, and so separate.

or (3) That the leaves at B will move toward the earthed casing in order to increase the capacity of the condenser formed by AB and the casing, and so reduce the potential energy $\frac{1}{2}\frac{Q^2}{C}$.

These explanations have been given when it was assumed that the charge C was positive. The explanation is the same, *mutatis mutandis*, in every particular, if it is realised in explanation (2) that negative charges move up the potential gradient, or against the field instead of with it.

Divergence of Leaves of Charged Electroscope

There is no difference whatever, from the point of view of the leaves of the electroscope, between a charged electroscope (whether charged directly or by induction) and an uncharged electroscope with opposite induced charges on leaves and knob, as in the last section. All three explanations of the divergence of the leaves therefore hold.

It is perhaps worth noting that no explanation has been offered for the stopping of the leaves in a definite position. In (1) and (2) we may suppose that they stop when the couple due to the electric force is exactly balanced by that due to the gravitational (and possibly elastic) forces on the leaves.

In (3) it is necessary to realise that the leaves can hold three kinds of potential energy—electrostatic, gravitational and elastic—and they will stop in equilibrium when the sum of these three kinds is a minimum.

If it be asked what happens to the energy lost when the position of least potential energy is found, we can only answer that it is dissipated irrecoverably as heat. If it were not for frictional and viscous resistances the leaves would oscillate indefinitely, but when they are damped down by these resistances their kinetic energy is converted into heat.

Charging an Electroscope by Induction

If when a charged body is near an uncharged electroscope (so that the induced charge on the leaves is causing them to diverge) the electroscope is earthed, the leaves will collapse

are earthed, preferably to a water-pipe through a soldered connection.

The other end of the H.T. battery is connected to a high resistance which may take the form of two wires in water in a glass **U**-tube (W in the figure). This forms a safety-precaution to guard the suspension from disaster in an accidental short-circuit. A high resistance, of the order of 100,000 ohms, will do equally well for W.

The needle of the electrometer is connected to a two-way switch S which can go either to the H.T. battery through the water-resistance or to earth.

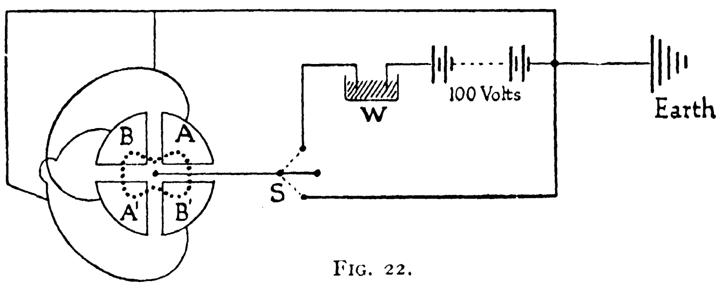


FIG. 22.

5. The needle is put to a potential of 100 volts with respect to the quadrants by means of the switch S.

It is probably deflected, but if the initial adjustments have been reasonably accurate it should not go off the scale.

The needle is then brought back to zero by adjusting the main levelling-screws. It is then earthed by the switch, and probably goes off the scale again, but not so far. It is again adjusted by the levelling-screws, and again charged.

The whole process is repeated until the zero reading is unchanged when the needle is charged, the quadrants remaining earthed.

6. The quadrants BB′ are now connected to a potentiometer circuit as in Fig. 23. S_1 is the original two-way switch. S_2 is a reversing-switch. V is a two-volt accumulator. R is a subdivided high resistance of as large a value as possible with a total resistance of not less than about 10,000 ohms, arranged so that connection can be made to any point on it,

and so that the ratio of the resistance from one end to the point of contact to the whole resistance is known with an accuracy of 1 in 1000.

A point of contact on R is found so that when the needle is charged the deflection is within the limits of the scale on both sides. The reversing-switch gives a quick change from one side to the other.

The deflection for any one point of contact should be the same on both sides. If it is not, the effect of small adjust-

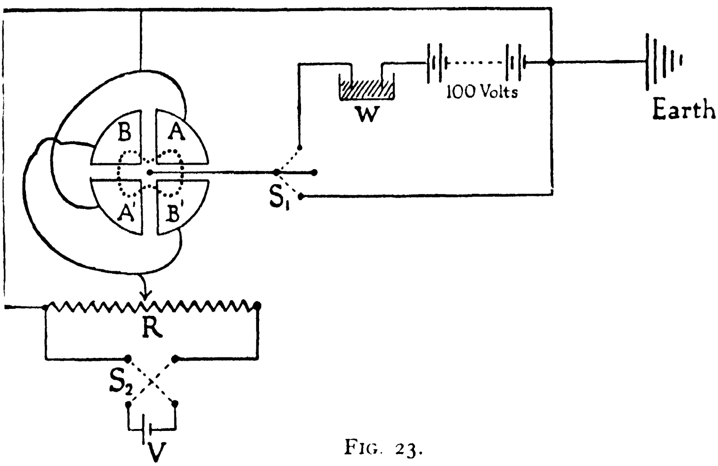


FIG. 23.

ments of the levelling-screws should be tried. These may disturb the zero again. The whole process of adjustment should now go on until the best possible compromise has been found. A zero error of 0·5 cm. and a percentage difference of 5% between the two deflections may be regarded as good enough with the scale at 100 cm. from the electrometer. Quadrant electrometers sometimes show signs of demoniac possession.

Experiments with the Quadrant Electrometer

Sensitivity to Potential.—With the circuit as in Fig. 23 values of the deflection on both sides are observed from deflections of about 1 cm. to deflections of about 20 cm. with the

scale at 100 cm. distance, R being varied. The deflection should be proportional to the potential difference between the quadrants. A graph of mean deflection against potential of quadrants BB′ is drawn, and the average deflection in cm. per volt is found. It is of the order of 20 cm. for suspensions of ordinary sensitiveness.

Variation of Sensitivity with Needle Potential.—The same circuit is used, but the 100-volts needle potential is increased to the largest value obtainable (or regarded as safe) in the laboratory, the potentiometer R being adjusted meanwhile so that both quadrants are earthed.

R is then adjusted so that the deflection is just not too great for the scale on either side.

R is now left fixed, and the deflection is then observed for needle-potentials from zero up to the maximum value. The result is plotted on a graph.

It will be found that this graph is not a straight line, but a curve convex upwards if needle-potential is plotted on the horizontal axis and deflection on the vertical axis.

The sensitivity should be found to be nearly proportional to the needle-potential for low values, but to fall off for higher values.

Since by the theory of the instrument (see p. 429) the deflection for a given potential difference between the quadrants should be proportional to $\frac{V}{1 + \alpha V^2}$, where V is the needle-potential, this result is to be expected.

The greatest value of $\frac{V}{1 + \alpha V^2}$ is when $V = \sqrt{\frac{1}{\alpha}}$.

We can easily find the value of V giving maximum sensitivity from our graph. For if we assume that the deflection is proportional to $\frac{V}{1 + \alpha V^2}$, then

$$d = \frac{KV}{1 + \alpha V^2}.$$

So $\alpha d V^2 - KV + d = 0$.

Any two corresponding sets of values of V and d will provide us with a pair of simultaneous equations giving K and α, and several pairs should give us a good mean value for α.

So since $V = \frac{1}{\sqrt{\alpha}}$ for maximum sensitivity, this value of V can be calculated. It is probably of the order of 500 volts.

Measurement of Alternating Potentials. Idiostatic Use of Electrometer.—In the first two experiments the electrometer was said to be used Heterostatically. If one of the quadrants is connected to the needle it is said to be used Idiostatically. The circuit is connected as in Fig. 24. V is a source of potential (either a set of cells or a potentiometer) giving potential differences up to about 12 volts for an electrometer giving 20 cm. deflection with 1 volt between the quadrants

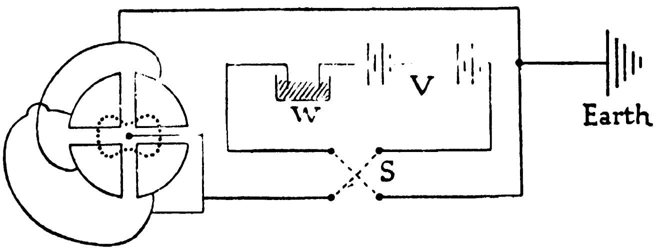


FIG. 24.

and 100 volts on the needle in the heterostatic arrangement. S is a reversing-switch.

Now we have, if v be the potential of one quadrant and V the needle-potential in the heterostatic arrangement, the deflection d given by

$$d = \frac{KVv}{1 + \alpha V^2}.$$

In the idiostatic arrangement, since $v = V$

$$d = \frac{KV^2}{1 + \alpha V^2}.$$

Since V only occurs in terms of the second order the deflection is in the same direction whatever the sign of V. Moreover, since comparatively small values of V are used, the αV^2 term in the denominator is practically negligible, and we have

$$d = KV^2.$$

d is now observed for various values of V, the sign of V

being changed by the reversing switch for each observation. d is then plotted against V^2, and an approximate straight line should be obtained.

If a sufficiently low-voltage source of alternating current can be obtained the mean value of the potential may be measured with a quadrant electrometer used idiostatically.

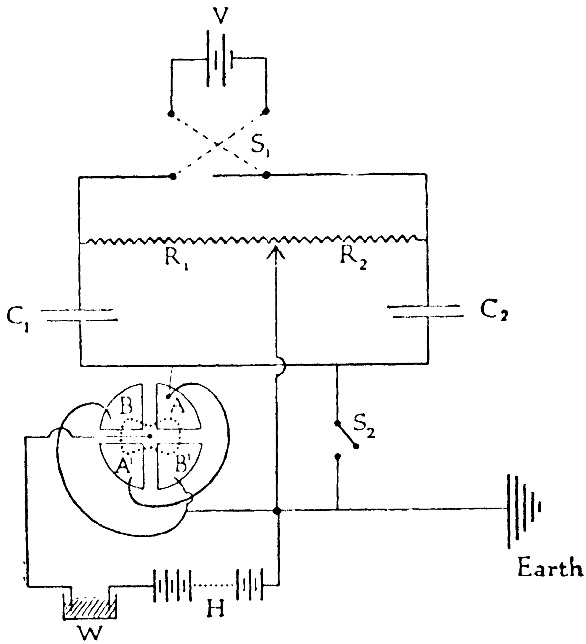


FIG. 25.

This mean value may be shown to be $\frac{1}{\sqrt{2}}$ of the maximum value if the potential v at time t is given by

$$V = V_0 \sin pt.$$

Where V_0 is the maximum value of V, and the frequency of the oscillations is $\frac{p}{2\pi}$, $\frac{1}{\sqrt{2}}$ can be shown to be the average value of sin pt. See p. 442.

Comparison of Capacities.—The circuit is connected as in Fig. 25.

V is an accumulator, H a high-tension battery (about 100 volts), W a high-resistance, S_1 a reversing-switch, S_2 an insulated on-and-off switch, C_1 and C_2 the condensers whose capacities are to be compared, and R_1 and R_2 the resistances of the two parts of a high-resistance potentiometer. The ratio of R_1 to R_2 must be accurately known.

If a current i be flowing through R_1R_2, then the potential of the top plate of C_1 will be R_1i, and that of C_2 will be $-R_2i$.

Thus the induced charges on the bottom plate of C_1 will be C_1R_1i, and that on C_2 will be $-C_2R_2i$.

The total charge on the conductor formed by connecting the lower plates of C_1 and C_2 and the quadrants A and A′ will then be

$$(C_1R_1i - C_2R_2i).$$

This will be zero if $C_1R_1 = C_2R_2$,

$$\text{that is, if } \frac{C_2}{C_1} = \frac{R_1}{R_2}.$$

Let the switch S_1 be open.

If the switch S_2 be closed and opened, the quadrants AA′ will be left insulated and uncharged.

If S_1 be now closed in either direction there will be no deflection if, and only if, no charge is given to the quadrants; that is, if $\frac{C_2}{C_1} = \frac{R_1}{R_2}$.

We have thus a method of comparing the capacities of two condensers. If we use S_2 as a reversing switch instead of simply as an on-and-off switch the sensitivity of the method is doubled.

The Attracted-disc Electrometer

The only instrument which can be used directly for the absolute determination of potential differences is the Attracted-Disc Electrometer, invented by Lord Kelvin. It is essentially a parallel-plate condenser, fitted with a guard-ring, so arranged that the force of attraction and the distance between the plates can be simply and accurately measured. Let A be the insulated plate, of area a sq. cm., and B the earthed

plate. A is usually held by an adjustable spring, and the load required to bring it level with its guard-ring plates, when A and B are at the same potential, is found (Fig. 26).

Actually A is held by a small spring. First the plates are both earthed and a small load m is put on A. This extends the spring and lowers A below its guard-ring. The column supporting the spring is now raised by a screw-motion till A is level with its guard-ring. The load m is now removed, and A rises above the guard-ring CD. It is then given potential V, and the level of the earthed plate B is altered until the electrostatic attraction between the plates brings A back to its zero. The electrostatic attraction is now equal to mg. It is difficult to tell with this electrometer, as with all arrangements involving a parallel-plate condenser, where the zero of t, the distance between the plates, really is. So usually two values of the potential V of the plate A are applied in turn, and t is observed for each of them.

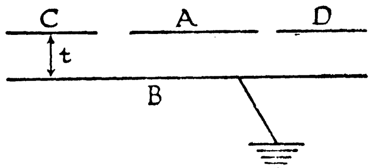


FIG. 26.

Suppose t is the distance for potential V. [This supposition is to make the calculation simple. We do not know the zero of t. But we do know that any two special values of t, t_1 and t_2 are measured from the same zero, so that a change of $(t_1 - t_2)$ would accurately correspond to a change $(V_1 - V_2)$.]

Let σ be the surface density of charge on the plate A. Then $mg = 2\pi\sigma^2 a$, since mechanical force per sq. cm. $= 2\pi\sigma^2$, as shown on p. 44, if a is the area of the plate A. Now since the field between the plates is uniform

$$\text{Electric Intensity} = 4\pi\sigma = \frac{V}{t}$$

$$\therefore \quad \sigma = \frac{V}{4\pi t}.$$

$$\therefore \quad mg = 2\pi a \cdot \frac{V^2}{16\pi^2 t^2}.$$

$$\therefore \quad V = t\sqrt{\frac{8\pi mg}{a}}.$$

If a distance t_1 corresponds to V_1 and t_2 corresponds to V_2

$$(V_1 - V_2) = (t_1 - t_2)\sqrt{\frac{8\pi mg}{a}}.$$

All the quantities on the right-hand side of this equation are directly measurable. So the value of $(V_1 - V_2)$ can be calculated in absolute units from first principles.

CHAPTER V

ELEMENTARY MAGNETISM

Elements of Magnetism—Elementary Experiments—Resemblances and Differences between Magnets and Charged Bodies.

Elements of Magnetism

As in Electrostatics, so in Magnetism, Thales of Miletus was the earliest known experimenter.

He is thought to have discovered the Lodestone, which came from Magnesia in Asia Minor, and so gave a name to the subject. Dr. Gilbert, however, was the first worker to leave a record of his work. This he did in his great book, *De Magnete Magneticisque Corporibus, et De Magno Magnete Tellure; Physiologa Nova.* ("Of the Magnet and Magnetic Bodies, and of the Great Magnet, the Earth; a New Discussion of their Nature.")

This book was published in 1600, and for more than 200 years a close relation between magnetism and electricity was suspected but never discovered. Not till 1819 did Oersted observe the magnetic field due to an electric current.

Before we come to units and exact definitions, we may summarize the elementary facts about magnetism.

A lodestone is found to orient itself (or point) in a particular direction when it is freely suspended. If it is put in iron filings the filings are found to cluster more thickly about two regions on the stone than anywhere else. When the stone is freely suspended the line joining these regions runs roughly north and south. The regions are given the name of "poles," and the one pointing north is called the North Pole of the lodestone, and the other the South Pole.

It is found that like poles of different lodestones repel one another and unlike poles attract one another. Thus it is inferred that the earth has a pole of South polarity some-

where near its geographical North Pole, and a pole of North polarity somewhere near its geographical South Pole.

If a piece of soft iron is put near a lodestone it is found to acquire the same properties as the lodestone so long as it remains near. It possesses poles, and is said to be magnetized by induction. If the piece of iron is near the North Pole of the lodestone the end nearest the lodestone is a South pole. If the iron is near the South pole of the lodestone the end nearest the lodestone is a North Pole. If a piece of hard steel is stroked from one end to the other by one of the poles of a lodestone it acquires the lodestone's properties permanently. The end of the steel where the stroking leaves off is found to have polarity opposite to that of the stroking pole. The end where the stroking begins is found to have polarity like that of the stroking pole.

If such a permanent magnet (or a lodestone) is cut in half across the line joining its poles (which is called the Magnetic Axis), each of the halves is found to be a complete magnet with poles. It is impossible by any means to get one pole without the other.

If a piece of steel is stroked from each end to the middle, or from the middle to each end, with a pole of another magnet it is found to be magnetized with like poles at the two ends and an opposite pole (or rather two of them together) in the middle. Such a magnet is said to have consequent poles.

The region round a magnet is known as a " Magnetic Field." If an unmagnetized piece of iron is held in a magnetic field, it is found to become a magnet. If it is hammered it becomes a stronger magnet; but if it is hammered when it is not in a magnetic field it tends to lose its magnetism.

If a bar of iron is held, at any place in England well away from other magnets or masses of iron, with its axis pointing North and downwards at an angle of about 67° to the horizontal, it becomes, when hammered, a stronger magnet than if held in any other position. If it is hammered when held at right angles to this position it tends to lose its magnetism.

A magnet quite freely suspended in England, away from all other magnetic materials, comes to rest with its axis pointing North and downwards at an angle of about 67° to the horizontal.

If a magnet is heated to red heat it loses its magnetic properties and does not recover them on cooling.

If magnets are placed under a sheet of paper, and iron filings are lightly sprinkled on top of the paper, they are found to arrange themselves in very striking patterns, following definite lines which seem always to go from one pole to another if possible. [The effect is shown best if the filings are blown off a tray on to the paper, and the paper is lightly tapped afterwards.]

If a small compass is placed on the paper it is always found to point in a direction parallel with the nearest line traced out by the filings. Such lines are called lines of force.

We are only dealing now with the very roughest ideas of magnetic poles, magnetic fields, and magnetic lines of force. Exact definitions which can be used mathematically will be introduced as soon as possible.

Elementary Experiments

It is vitally important that everyone learning magnetism should perform the following (or an equivalent) series of experiments before going on to quantitative work.

(1) Take a bit of lodestone and verify the fact that it attracts iron filings; if you are lucky, you may succeed in showing that it orients itself in a particular direction if you hang it up or set it floating on a piece of wood. Ordinary cotton tends to untwist if any strain is put upon it, so you must either get a single strand of cotton or try to allow for the untwisting, which is a very difficult and unsatisfactory thing to do.

(2) Hang a strong bar magnet in a paper stirrup from a wooden stand, and observe that it orients itself approximately north and south if you can circumvent the twist of the suspension.

(3) Verify that the north pole of another bar magnet repels the north pole of the suspended one, that the south poles similarly repel one another, and that opposite poles attract one another.

(4) Test a set of new knitting-needles with a pocket compass till you find one which is completely unmagnetized. Hang it up on the finest possible suspension, and verify that it does not set in any particular direction. Magnetize it by stroking it with the bar magnet, and verify that it now orients itself in the expected direction.

(5) Magnetize a sealed glass tube containing iron filings, and verify with a compass that it has poles. Watch carefully the behaviour of the filings as the magnet passes them. It illustrates the supposed behaviour of soft iron near a magnet.

(6) Magnetize a soft iron rod (preferably about two inches long). Make it red-hot in a Bunsen flame. Cool it and test it to see if it is demagnetized.

(7) Magnetize a large soft iron rod. Lay it on the floor pointing east and west and hit it several times with a hammer. Test it to see if its magnetism is destroyed or considerably reduced.

(8) Take the same rod and hold it in a vertical plane running north and south, and point it down toward the north so that it makes an angle of about 67° with the horizontal. [This is the true direction of the earth's field in England. It will be explained in more detail in Chapter VII.]

Hit the rod several times with a hammer while you are holding it in this position, and then test it with the small compass for magnetization. It should be magnetized with the north-pointing end a north pole.

(9) Pick up a piece of soft iron by a bar magnet and hang it in iron filings. The soft iron should behave like a fairly strong magnet, so long as it is kept magnetized by the presence of the bar magnet; but when the bar magnet is removed the soft iron should lose nearly all its magnetism, and the filings should almost all drop off.

(10) Fix an unmagnetized soft iron rod horizontally in a wooden clamp on a wooden stand. Fix a compass near one end of it. The compass may be slightly, but should not be considerably, affected by it.

Bring one pole of a bar magnet near the other end of the soft iron rod. The latter should now become a strong magnet and produce a very marked effect on the compass.

You may think this effect is mainly due to the bar magnet, since some of it certainly is. Remove the soft iron rod without moving either the bar magnet or the compass, and the effect on the compass should now be much less; so that the original effect was evidently due mainly to induced magnetism in the soft iron rod, rather like the induced charge on an electroscope, negative on the leaves and positive on the knob, when a charged ebonite rod was brought near it.

(11) If the electroscope had had its knob chopped off by an insulated chopper while it was holding its charge, the knob would have kept a positive charge and the leaf a negative charge; but

no chopping of the iron rod with induced magnetic poles on it could have separated the poles.

Verify this by cutting a magnetized knitting-needle with a file into as many pieces as your time and patience allow, and test the magnetism of each of them with a small compass. Each should be found to have north and south poles, and the direction of magnetism of each piece should be unchanged.

(12) Take an unmagnetized knitting-needle, and magnetize it by stroking it from the ends to the middle with the north pole of a bar magnet. When you test it for magnetism, it should be found to have a north pole at each end, and a south pole (or rather two south poles very close together), in the middle. Such a magnet—a freak—is said to have "consequent poles."

(13) Get a strongly magnetized bit of knitting-needle, or a short, thin bar magnet, and stick it through a flat cork large enough to float when holding it.

Set it afloat in a trough with one of its poles just above the top of the cork.

Fix a big horseshoe magnet at the edge of the trough, with its poles facing into the trough. Bring the floating magnet near the horseshoe magnet so that its upper pole is near the like (repelling) pole of the horseshoe magnet.

It should be possible to make the floating magnet go to the unlike pole of the horseshoe magnet, not in a straight line but by following the curve of the field, whose general shape you should already realize from having plotted fields with iron filings.

This should show that what influences the movement of a pole should be regarded as the field near that pole, rather than the attraction or repulsion of comparatively distant poles.

This is of no immediate importance for elementary examinations, but is of first-class importance in the whole of physics, wherever the idea of attraction comes in, whether in connection with magnetic poles, electric charges, or gravitation.

(14) Plot the fields due to as many arrangements of two bar magnets, the horseshoe magnet, and the soft iron as you have patience for. Do the plotting by putting a large sheet of paper over your magnets, sprinkling filings on them, and giving the paper some light taps.

Sketch each arrangement.

Apparatus. Lodestone, iron filings, large sheet of paper, water-trough, flat piece of wood, thin cotton, wooden retort-stand, two strong bar magnets, set of new knitting-needles, pocket compass, test-tube with cork, soft iron rod, hammer, file, small flat cork, large horseshoe magnet.

Resemblances and Differences between Magnets and Charged Bodies

It is interesting to consider the puzzle which was presented to the early workers during the two hundred years when they knew the elementary facts of Electricity and Magnetism, but had not related the two by means of Electro-magnetism.

1. A magnet attracts pieces of magnetic material only; a charged body attracts any light object.

2. A magnet orients itself in a particular direction if it is freely suspended; a charged body does not.

3. A magnet has polarity, for the magnetism of its North pole is in some sense opposite to that of its South pole; a body with an induced charge has this property, being positive at one end and negative at the other; but a body with a free charge has not.

5. A magnet makes a piece of magnetic material in its neighbourhood into an induced magnet, and the induced magnetism disappears (for soft iron) when the magnet is removed; a charged body has exactly corresponding properties, but will induce a charge on any insulated conductor instead of on particular ones only.

6. Touching with the hand has no effect on a magnet; but it discharges a charged body.

7. Cutting a magnet in half gives us two magnets; cutting a body with an induced charge in half gives us two bodies with opposite charges (if our knife and the charged body are both insulated).

8. All magnets contain both kinds of magnetism; charged bodies (save in the case of induced charges) have one charge only.

9. If we bring two magnets together, nothing happens except that they exert a force on each other; if we bring two oppositely charged bodies together a spark passes between them.

10. If we have a pair of long enough ball-ended magnets we can show directly that their poles exert forces on each other in a manner consistent with the Inverse Square Law; we can show the same effect for charged bodies.

This last fact was established by Coulomb, who performed

his Torsion-Balance experiment for individual poles of long magnets as well as for electric charges, and found his results consistent with the Inverse Square Law within the limits of his experimental error. Magnetic Lines of Force[1] resemble Electric Lines of Force in their general behaviour, though there are important differences.

[1] It is difficult to summarise their properties at an early stage The following references may however be useful.

Magnetic lines of force possess properties (1), (3), (6), and (7) of those of electric lines of force given on pp 28 and 29, except that the word "magnetic" should be substituted for the word "electric" in properties (1) and (7).

The unit of magnetic intensity—the Oersted—referred to in properties (1) and (7) is defined on p. 72.

Magnetic potential, which would apply to property (3), p. 29, is considered on p. 330.

Properties (2), (4), (8), and (9) on p. 29 have no useful close analogues for magnetic lines.

The further properties of magnetic lines of force may perhaps be understood from a study of the diagrams on pp. 218–229, 359–360, and (taken together) 238, 241 and 357, and the text associated with such diagrams.

Lines of magnetic induction are discussed on pp. 358–360, and compared with tubes of electric induction on pp. 31–32.

It is advisable for the student at a later stage to write an essay on Magnetic Lines of Force and Induction.

CHAPTER VI

QUANTITATIVE MAGNETISM—POLES AND FIELDS

Magnetic Poles—Magnetic Intensity—Permeability—The Deflection Magnetometer—Magnetic Moment—Magnetic Intensity on the Axis of a Bar-magnet—Magnetic Intensity on the Perpendicular Bisector of the Axis of a Bar-magnet—Gauss's Proof of the Inverse Square Law—End-on and Broadside Positions—Neutral Points—Oscillations of a Suspended Magnet—Determination of M and H from First Principles—The Vibration Magnetometer—Energy Stored in a Deflected Magnet.

Magnetic Poles

It is very easy to find in any elementary text-book a rough definition of a magnetic pole. Such a definition was given in the last chapter. It is also easy to find an exact definition of a Unit Magnetic Pole. It is a magnetic pole which exerts a force of 1 dyne on an equal pole 1 cm. away *in vacuo*, when each pole is concentrated at a point. But it is exceedingly difficult to find a really satisfying definition (or explanation) of what the term magnetic pole really means when it is used in connection with strength rather than position. If you consider the elementary definition in the last chapter you will see that it really depends on position only. This sort of problem—the problem of knowing exactly what one is talking about—is so important in Physics that it is worth while investigating it carefully.

Let us suppose that an intelligent foreigner from a far country where butter is unknown is suddenly sent out, on the first day of his first visit to England, to buy a pound of butter. He can speak English, but is too shy to ask the housewife who sends him what she means by a pound of butter. She is cross that morning, and as his first desire is for peace, he hurries out into the street, without stopping to consider that he does not know any more than the name of what he is required to get.

In the street he meets a friend, and asks the friend to explain what "a pound of butter" means. The friend explains that a pound of butter is a lump about 6″ by $2\frac{1}{4}$″ by $2\frac{1}{4}$″, which balances a brass weight labelled 1 lb. when they are put in the scales. This is true, but is no use to the foreigner. It defines his unit quantitatively, but gives him no idea of its nature. He asks what this butter is, and learns that it is a yellow product of the milk of the cow. His own country contains no cows. Ultimately the only solution of the problem is to show him some butter and let him handle and taste it. The word "butter" then at last has a definite meaning for him, and he knows that what he means is what everyone else means by the word. Unless he is a chemist he still cannot define it, though he can recognize it, and is satisfied about its meaning.

We speak of "quantity of pole" or "strength of pole" just as we speak of quantity of charge. When we do this we imagine a magnet to be permeated by a magnetic material which is particularly concentrated in two or more definite regions of the magnet. This magnetic material, like electric charge, may be of two opposite kinds called North and South, or Positive and Negative. The regions where it is particularly concentrated are loosely called the "poles" of the magnet, and the "strength" of a particular pole depends on the quantity of the material in its neighbourhood.

It is important to understand that this magnetic material, unlike electric charge, is entirely imaginary and can never be isolated. We can only *observe* magnetic fields, and we *infer* magnetic poles as a mathematical convenience to account for the fields. Electric charges exist, but magnetic poles do not.

We can now give a workable definition of Magnetic Pole and of Unit Magnetic Pole.

"A Magnetic Pole is a quantity of an imaginary magnetic material which is supposed to permeate a magnetized body, and which may be represented as concentrated at a point."

"A Unit Magnetic Pole is of such strength that it exerts a force of 1 dyne on an equal pole 1 cm. away *in vacuo*, each pole being concentrated at a point. It is sometimes called a Weber."

Magnetic Intensity

As magnetic pole corresponds to electric charge, so does Magnetic Intensity correspond to Electric Intensity. The magnetic intensity at any point is a vector which determines the force on a magnetic pole situated at the point. The numerical value of the magnetic intensity is the same as that of the force in dynes on a unit pole, and its direction is that of the force on a north pole.

We may now define unit magnetic intensity.

"Unit Magnetic Intensity is the intensity *in vacuo* at a distance of 1 cm. from a unit point-pole; and also the intensity which causes a unit pole to experience a force of 1 dyne. It is called an oersted. Its direction is the direction of the force on a north pole."

Coulomb's experiment pointed (with magnetic poles as with electric charges) to the probability that the force between two magnetic poles *in vacuo* was proportional to the product of their pole-strengths and to the inverse square of the distance between them. Nothing has since been discovered to make us question this assumption in the problems in which it is useful, and we therefore accept it.

It thus follows from our choice of units that, if two poles of strength m_1 and m_2 webers are r cm. apart *in vacuo*, the force F between them in dynes is given by

$$F = \frac{m_1 m_2}{r^2}.$$

Also it follows that the magnetic intensity H oersted at a distance of r cm. from a pole of strength m webers *in vacuo* (no other pole being near) is given by

$$H = \frac{m}{r^2}.$$

It also follows that the force F in dynes on a pole of strength m webers in a field of intensity H oersted is given by

$$F = Hm$$

Permeability

If a uniform material medium occupies the space between the poles the force between them is found to be changed.

The ratio $\dfrac{\text{Force in vacuo}}{\text{Force in medium}}$ is known as μ, the Permeability. It thus corresponds to k, the dielectric constant for media between electric charges.

When a medium of permeability μ is present the first two equations above thus become

$$F = \frac{m_1 m_2}{\mu r^2}$$
$$H = \frac{m}{\mu r^2},$$

but the third is, of course, unchanged, since the force F on m_2 in the first equation may be regarded as $\left(\dfrac{m_1}{\mu r^2}\right) \times m_2$, or Hm_2.

When μ is much greater than 1, the medium is called Ferromagnetic. When μ is a very little greater than 1 the medium is Paramagnetic. When μ is a very little less than 1 the medium is Diamagnetic. μ is never much less than 1. These terms are dealt with in more detail in Chap. V, Part II.

The Deflection Magnetometer

The Deflection Magnetometer is an instrument for comparing two magnetic intensities which are acting at right angles to one another. One of these intensities is usually in practice the horizontal component of the earth's magnetic intensity.

A Deflection Magnetometer consists of a very small magnet able to turn very freely in a horizontal plane. It has a light double-ended non-magnetic pointer fixed at right angles to it. This pointer has a mirror and a scale below it, the scale reading in degrees from zero to 90 in both directions at both ends. The purpose of the mirror is to prevent parallax error in reading the scale. If one sees the image of the pointer in the mirror obscured by the pointer when the reading is taken, one must be looking vertically downwards. Both ends of the pointer are read, and the mean is taken.

It can be shown as follows that if an intensity F is acting at right angles horizontally to the horizontal component H

of the earth's magnetic field, then if θ is the angle through which the pointer is deflected

$$F = H \tan \theta.$$

Let us now consider with a purely geometrical figure exactly what happens.

Suppose NS (Fig. 27) is the central line of the magnet (not shown because it would take up too much room). Let O be the centre of the magnet, on which it is pivoted, and let

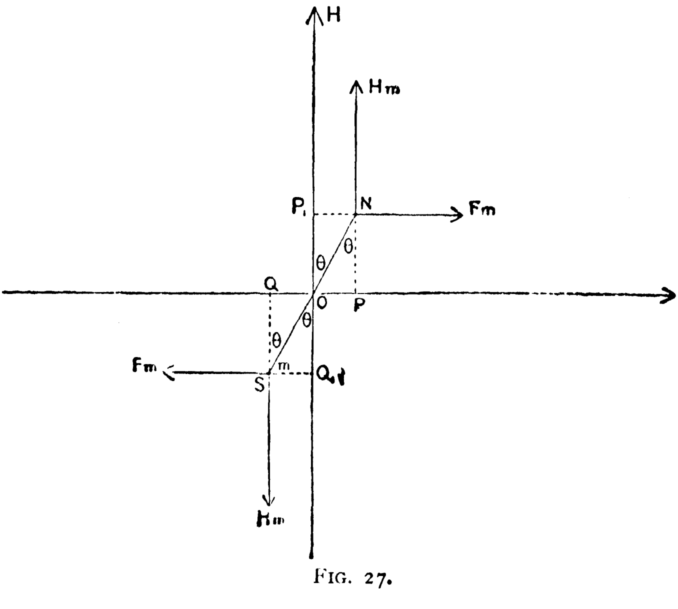


FIG. 27.

NP and NP_1 be perpendiculars dropped from N on the lines of action of F and H respectively. Let SQ and SQ_1 be corresponding perpendiculars from S. Let the pole-strength of the magnet be m.

The geometry of the figure shows that

$$\angle PNO = \angle NOP_1 = \angle QSO = \angle SOQ_1 = \theta.$$

The field H exerts a force of Hm dynes in the direction of H on the pole N of pole-strength m, and an equal force in the opposite direction on the pole S.

The field F exerts a force of Fm dynes in the direction of F on the pole N, and an equal force in the opposite direction on the pole S.

The moment about O, due to the force H on N, is equal to OP.Hm, and

$$\begin{aligned}\text{Moment due to F on N} &= OP_1.Fm.\\ \text{Moment due to H on S} &= OQ.Hm.\\ \text{Moment due to F on S} &= OQ_1.Fm.\end{aligned}$$

These four moments are the only ones trying to turn the magnet round. The moments due to F act in a clockwise direction, and those due to H act in an anti-clockwise direction. When the magnet is still, these two moments clearly are equal and opposite. Then

$$\text{Moment due to F} = \text{Moment due to H.}$$

$$\therefore\ OP_1.F.m. + OQ_1.F.m. = OP.H.m. + OQ.H.m.$$

But, from the geometry of the figure, $OQ_1 = OP_1$, and $OQ = OP$.

$$\therefore\ 2\ OP_1.m.F = 2OP.m.H.$$

$$\therefore\ F = \frac{OP}{OP_1}.H.$$

$$\therefore\ F = H \tan \theta.$$

Magnetic Moment

In practice we never (except in very rough approximations) consider the intensity due to a single magnetic pole, since magnetic poles always occur in pairs. A more useful constant of a magnet than its pole-strength is its Magnetic Moment. It is the first directly measurable quantity met with in Magnetism. The best definition of Magnetic Moment is as follows :

" The Magnetic Moment of a Magnet is the moment of the couple acting on it when its axis is perpendicular to a magnetic field of unit strength."

This is an excellent definition, because it has no special requirements about the distribution of pole in the magnet. It is not, however, easy to deal with in elementary equations. For these a definition which is mathematically simpler, but less in touch with reality, is used.

Let us suppose that our magnet has point-poles, each of strength m webers, situated $2l$ cm. apart.

The force on each of these poles in a field of strength 1 gauss is m dynes. If the line joining the poles is perpendicular to the field the two forces on the poles form a couple. The forces of the couple are m dynes each, and they are $2l$ cm. apart. So the moment of the couple is $2ml$ dynes.

But $2ml$ is the product of the pole-strength and the distance between the poles. We thus have our second definition :

"The Magnetic Moment of a Magnet is the product of the pole-strength and the distance between the poles."

This definition assumes that the poles are concentrated at points, which is never true.

The assumption that the poles are concentrated at points is, however, distinctly useful, for either the pole-strength of a bar-magnet is determined so as to make it fit this definition after the length has been fixed, or the length is determined to fit after the pole-strength has been fixed.

The length may be determined by choosing arbitrarily what points are most nearly aimed at by the general average of the lines of force entering the magnet. Iron filings give the approximate distribution of the lines of force. The distance between these points is $2l$. The magnetic moment M is determined independently, and the pole-strength is then said to be given by

$$m = \frac{M}{2l}.$$

Alternatively the magnet may be pulled suddenly through a coil in series with a ballistic galvanometer, from a position where the plane of the coil bisects the magnet approximately. The throw of the galvanometer gives the number of lines of force leaving each end of the magnet, if the galvanometer has been calibrated for charge. The mean of the throws for the two ends gives a direct measure of the number of units of pole in each half of the magnet. This gives a genuine value of the pole-strength m. The distance between the poles is then said to be given by

$$2l = \frac{M}{m}.$$

So all goes well in practice. If, however, we determined M by the couple directly, m by the ballistic galvanometer, and $2l$ by direct observation, it is unlikely that the product of m and $2l$ would be actually equal to M.

Magnetic Intensity on the Axis of a Bar-magnet

Let a point P on the axis of a bar-magnet be distant d cm. from the point O midway between the poles N and S each of strength m webers. Let NS $= 2l$ cm. (see Fig. 28).

Let us consider the intensity at P due to the magnet.

The intensity at P due to the near pole opposes that due to the far pole.

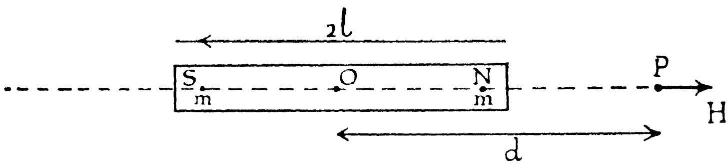


FIG. 28.

We have then that the intensity H at P is given by

$$\begin{aligned} H &= \frac{m}{PN^2} - \frac{m}{PS^2} \\ &= m\left\{\frac{1}{(d-l)^2} - \frac{1}{(d+l)^2}\right\} \\ &= \frac{m}{(d^2-l^2)^2}\{(d^2+2dl+l^2)-(d^2-2dl+l^2)\} \\ &= \frac{4mdl}{(d^2-l^2)^2}. \end{aligned}$$

Since $M = 2ml$, we may also write

$$H = \frac{2Md}{(d^2-l^2)^2}.$$

If the magnet is short, so that l^2 is negligible compared with d^2, we may put d^4 for $(d^2 - l^2)^2$ and so for short magnets

$$H = \frac{2M}{d^3}.$$

Magnetic Intensity on the Perpendicular Bisector of the Axis of a Bar-magnet

In this case (Fig. 29) the intensity due to N at P acts in the direction NP away from N, and that due to S acts along PS towards S; and these intensities are each equal to

$$\frac{m}{(d^2 + l^2)}$$

as may be seen from the figure, since $PN^2 = PS^2 = (d^2 + l^2)$.

The components of these intensities along OP are equal and

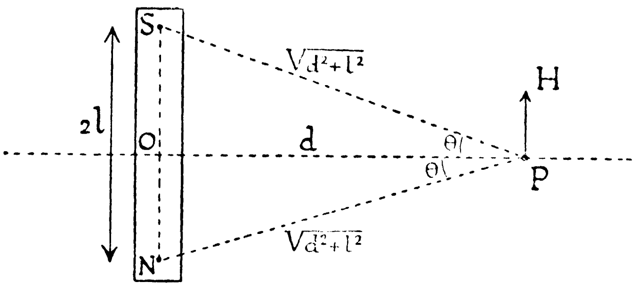


FIG. 29.

opposite, and therefore cancel each other. Their components perpendicular to OP reinforce each other. Thus their combined intensity acts in the direction of the arrow, and is equal to

$$\frac{2m}{d^2 + l^2} \cdot \sin\theta.$$

But $\sin\theta$ from the figure is $\dfrac{l}{\sqrt{d^2 + l^2}}$.

So
$$H = \frac{2ml}{(d^2 + l^2)^{\frac{3}{2}}}.$$

This may also be written

$$H = \frac{M}{(d^2 + l^2)^{\frac{3}{2}}}.$$

If l^2 is negligible compared with d^3,

$$H = \frac{M}{d^3}.$$

Gauss's Proof of the Inverse Square Law

Gauss gave a mathematical proof that if the law of force between magnetic poles were of the inverse nth power, then the intensity at any point on the axis of a short magnet is n times the intensity at a point the same distance away on the perpendicular bisector of the axis. Our results just obtained show that this holds for the inverse square law.

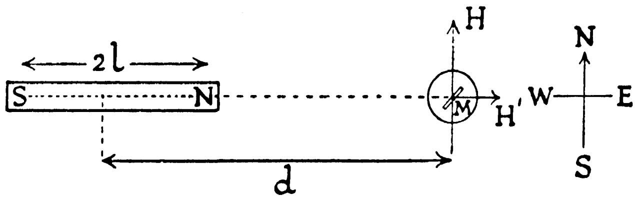


FIG. 30.

Gauss's experimental work showed that within the limits of error of his experiment the intensity on the axis was twice that at the same distance on the perpendicular bisector, and he therefore concluded that the law of force was that of the inverse square.

End-on and Broadside Positions

Two arrangements of a magnet with respect to a deflection Magnetometer cause the magnet's intensity to be perpendicular to the horizontal component of the earth's field. In both of these the magnet points along the magnetic east-and-west line.

In one, called the "End-on" or "A" position, the magnet's mid-point is on the same magnetic east-and-west lines as the magnetometer, and its axis is pointing at the magnetometer (Fig. 30).

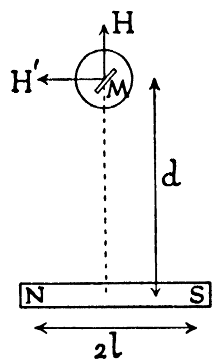


FIG. 31.

In the other, called the "Broadside" or "B" position, the magnet's mid-point is on the same magnetic north-and-south line as the magnetometer (Fig. 31).

In both figures, H shows the direction of the horizontal

component of the earth's field, and H′ the direction of the magnetic intensity due to the magnet.

In both positions, if the deflection is θ, $H' = H \tan \theta$.

In the A position, therefore

$$\frac{2Md}{(d^2 - l^2)^2} = H \tan \theta.$$

So
$$\frac{M}{H} = \frac{(d^2 - l^2)^2 \tan \theta}{2d}.$$

(a)

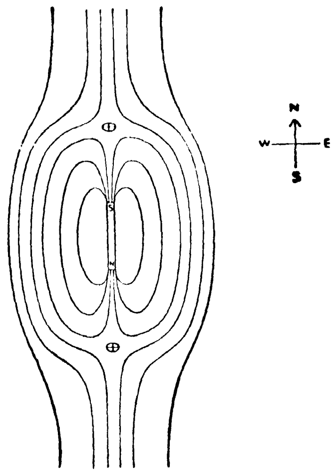


FIG. 32.

In the B position

$$\frac{M}{(d^2 + l^2)^{\frac{3}{2}}} = H \tan \theta.$$

So
$$\frac{M}{H} = (d^2 + l^2)^{\frac{3}{2}} \tan \theta.$$

The reason for the names " End-on " and " Broadside " is obvious.

Neutral Points

If a bar-magnet is placed with its axis along the magnetic north-and-south line, there are two points on its axis, as shown in Fig. 32, where H and H′ are equal and opposite and

the total intensity is zero. These points are called " neutral points." They are best found by plotting the lines of force round a magnet with one or more small exploring compasses.

If such points are distant d from the centre of the magnet it is clear that, since $H = H'$,

$$\frac{2Md}{(d^2 - l^2)^2} = H$$

so that

$$\frac{M}{H} = \frac{(d^2 - l^2)^2}{2d}.$$

If the magnet lies with its axis on the magnetic north-and-south line with its N pole pointing north, the neutral points

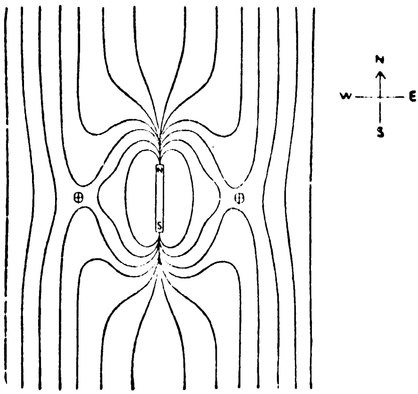


FIG. 33.

are clearly on the perpendicular bisector of the axis, and the aspect of the field is as in Fig. 33. We have then

$$\frac{M}{(d^2 + l^2)^{\frac{3}{2}}} = H$$

so that

$$\frac{M}{H} = (d^2 + l^2)^{\frac{3}{2}}.$$

Oscillations of a Suspended Magnet

If a magnet is suspended so that it can swing freely in a magnetic field H, its equilibrium position is obviously that in which its axis points in the direction of the intensity. If it is

displaced from this position through an angle θ, a restoring couple MH sin θ acts upon it. It therefore oscillates about its mean position as a pendulum does. The period of its oscillation can be shown to be given for small oscillations, in which we may take $\theta = \sin\theta$, by

$$T = 2\pi\sqrt{\frac{I}{MH}}$$

if I is the moment of inertia of the magnet.

The restoring couple on the magnet is clearly MHθ for a small oscillation. Since this is proportional to the angular displacement θ, the motion is simple harmonic, and the equation is that deduced in elementary text-books on dynamics.

If a is the length of the magnet and b its horizontal width, its cross-section being rectangular, it can be shown that

$$I = \frac{W(a^2 + b^2)}{12}$$

where W is its mass.

For a cylindrical bar of length a and radius r, $I = W\left(\frac{a^2}{12} + \frac{r^2}{4}\right)$. This equation is deduced in elementary text-books on dynamics.

By observing the average time T of a number of oscillations we can find the value of MH, for from the equation given we see that

$$MH = \frac{4\pi^2 I}{T^2}.$$

Determination of M and H from First Principles

We have already, by means of the Deflection Magnetometer and the method of Neutral points, several independent methods of finding $\frac{M}{H}$.

The oscillations of our magnet give us MH.

Since $\left(\frac{M}{H} \times MH\right) = M^2$ we can determine M without making

any assumptions about its magnitude. Since $\left(MH \div \frac{M}{H}\right) = H^2$ we can also determine H directly.

Though this is a long and clumsy way of measuring H, it is the way requiring least knowledge and least apparatus.

The Vibration Magnetometer

A Vibration Magnetometer is simply a small freely suspended light magnet, the frequency of whose oscillations is used to compare magnetic intensities. The torsional effect of the suspension must be so small as to be quite negligible.

If our magnet takes a time T_1 to swing in a field of intensity H_1, and T_2 for intensity H_2, then

$$\frac{T_1}{T_2} = \frac{2\pi\sqrt{\frac{I}{MH_1}}}{2\pi\sqrt{\frac{I}{MH_2}}}.$$

Since $2\pi\sqrt{\frac{I}{M}}$ cancels, we get

$$\frac{H_2}{H_1} = \frac{T_1^2}{T_2^2}.$$

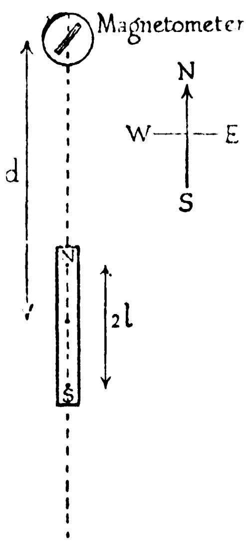


FIG. 34.

The use of the Vibration Magnetometer gives us the quickest way of comparing the absolute magnetic intensities at two different points.

If, however, we are dealing with the intensity of the field of a magnet, we must remember that the Vibration Magnetometer deals with the resultant of the magnet's field and the earth's.

If we want to compare the intensities of the fields of two magnets we must disentangle H from H_1 and H_2. In order to do this we must always have the magnet with its axis pointing north-and-south (not east-and-west as with the Deflection Magnetometer). The Vibration Magnetometer

may be suspended at any point either on the axis of the magnet (as in Fig. 34) or on its perpendicular bisector (as in Fig. 35).

Let us suppose that, as in the figures, the intensities H_1 and H_2 of the fields of the two magnets we are comparing are acting in the same direction as H. Then if T_0 is the time of

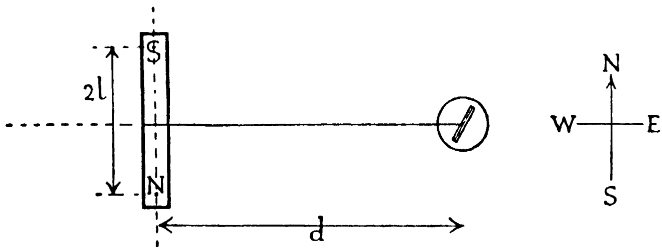


Fig. 35.

swing of the Vibration Magnetometer in H alone, T_1 in $(H_1 + H)$, and T_2 in $(H_2 + H)$, then

$$\frac{H_1 + H}{H} = \frac{T_0^2}{T_1^2}$$

so that
$$H_1 = \frac{(T_0^2 - T_1^2) \cdot H}{T_1^2}$$

Similarly
$$\frac{H_2 + H}{H} = \frac{T_0^2}{T_2^2}$$

so that
$$H_2 = \frac{(T_0^2 - T_2^2) \cdot H}{T_2^2}.$$

Then we have
$$\frac{H_2}{H_1} = \frac{\dfrac{(T_0^2 - T_2^2)H}{T_2^2}}{\dfrac{(T_0^2 - T_1^2)H}{T_1^2}}$$
$$= \frac{T_1^2(T_0^2 - T_2^2)}{T_2^2(T_0^2 - T_1^2)}.$$

This is rather a clumsy way of comparing the two intensities, but it is sometimes useful as a check. The Vibration Magnetometer may, however, be used to give us yet one more value of $\frac{M}{H}$.

Suppose H is obtained from the position of Fig. 35 with

the magnetometer on the perpendicular bisector of the magnet's axis.

Then $$H_1 = \frac{M}{(d^2 + l^2)^{\frac{3}{2}}}.$$

But also we have

$$H_1 = \frac{(T_0^2 - T_1^2) \, . \, H}{T_1^2}.$$

So $$\frac{M}{(d^2 + l^2)^{\frac{3}{2}}} = \frac{(T_0^2 - T_1^2)H}{T_1^2}.$$

So $$\frac{M}{H} = \frac{(T_0^2 - T_1^2)(d^2 + l^2)^{\frac{3}{2}}}{T_1^2}.$$

Energy Stored in a Deflected Magnet

If a magnet of length 2l and pole-strength m is turned through an angle θ in a magnetic field H, the distance moved by each pole in the direction of the field against the force of the field is obviously $(l - l \cos \theta)$. The force against which each pole is moved is mH. Thus the work done to turn the magnet to this position is

$$2mlH(1 - \cos \theta),$$

for no work is needed to move the pole across the field, since no force opposes its motion. Thus, if M is the magnetic moment, the work done in turning the magnet through an angle θ in a field H is

$$MH(1 - \cos \theta).$$

As the work done is equal to the potential energy stored, this is also the potential energy of the magnet.

As the magnet, when released, moves back, this work is changed to kinetic energy, so that, when the magnet is passing its position of equilibrium, with its axis along the field, its kinetic energy has the value which the potential energy had. As it moves on to the extreme position on the other side the kinetic energy is turned to potential again. The process goes on till all the energy is dissipated into stray heat by friction.

CHAPTER VII

TERRESTRIAL MAGNETISM

The Magnetic Elements—Magnetic Maps—The Secular Magnetic Variation—Periodic Magnetic Variations—Measurement of H—Measurement of Declination—Measurement of Dip—Causes of Terrestrial Magnetism—Possibility of Permanent Magnetism—Possibility of Earth-currents—Possibility of Currents in or above the Upper Atmosphere—The Sun's Magnetic Field—Further Suggestions of Causes—Lunar Magnetic Variations—Solar Magnetic Variations due to Radiation—Magnetic Variations due to Sunspots—Other Variations—The Kelvin Compass—Ships' Errors and their Correction. The Quadrantal Error—The Semicircular Error—The Heeling Error.

The Magnetic Elements

THE earth's magnetic intensity at any point on its surface is completely defined for strength and direction by three quantities, which together are known as the Magnetic Elements. These three are H, the horizontal component of the intensity; α, the Declination, the angle between the direction of H and the direction of true north; and θ, the Dip, the angle between the true direction of the intensity and the horizontal. Other quantities which may be required, but which are not included in the magnetic elements because they can be deduced when these are known, are I, the total intensity, and V, the vertical component of the intensity. It is obvious from Fig. 36 that

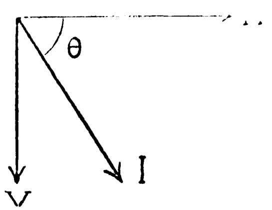


FIG. 36.

$$I = H \sec \theta$$

and

$$V = H \tan \theta.$$

α, the Declination, should more strictly be defined as the angle between the Geographical Meridian and the Magnetic Meridian.

At any point on the earth's surface the Geographical Meridian is a plane containing the point and the earth's North and

South Geographical Poles. This plane is necessarily vertical. The Magnetic Meridian is a vertical plane containing H (or I). It is important to notice that the Magnetic Meridian does *not* in general contain the earth's Magnetic Poles.

Magnetic Maps

The general arrangement of the earth's magnetic field is very roughly what it would be if the earth contained a bar

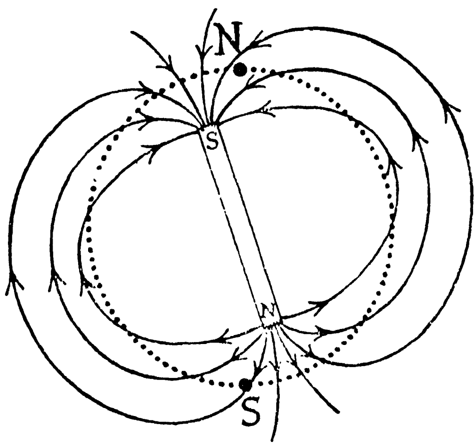


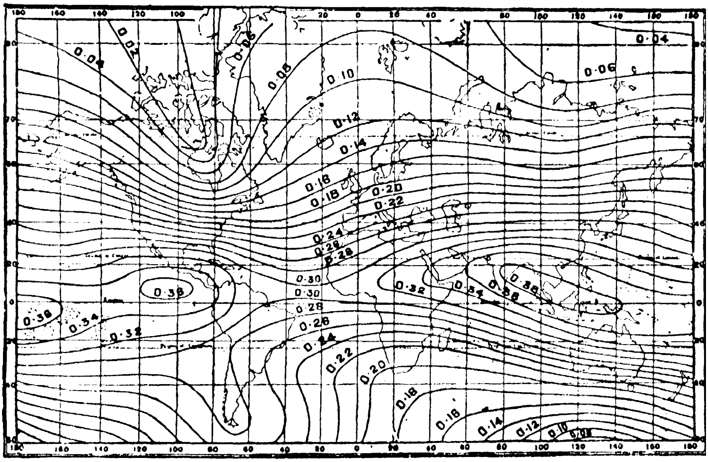
FIG. 37.

magnet with its south pole under Lat. 70° N., Long. 97° W. (Northern Canada) and its north pole under Lat. 72° S., Long. 154° E. (South Antarctic continent). Round an approximate great circle which is equatorial to these points I is approximately horizontal, and as one approaches either pole from this great circle I increases, but as θ also increases from 0° to 90° H diminishes till it is zero over the poles.

Fig. 37 gives a very rough diagram of the arrangement of the field; but it must not be supposed that the drawing of a large bar-magnet inside the earth has any significance whatever except that it produces a field of the right sort of general shape!

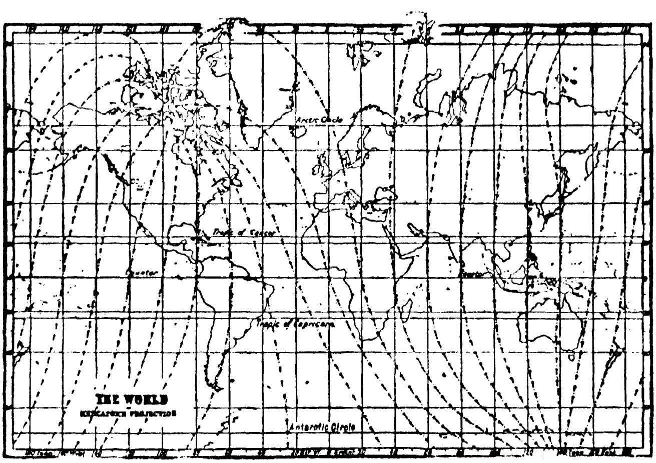
Many kinds of lines can be drawn on magnetic maps, which are used to represent various aspects of the earth's field.

Isogonals are lines of equal declination.
Agonic lines are lines of zero declination.



George Philip & Sons, Ltd. FIG. 38. London Geographical Institute

Isoclinics are lines of equal dip.
The *aclinic line* is a line of zero dip. It is also called the magnetic equator.

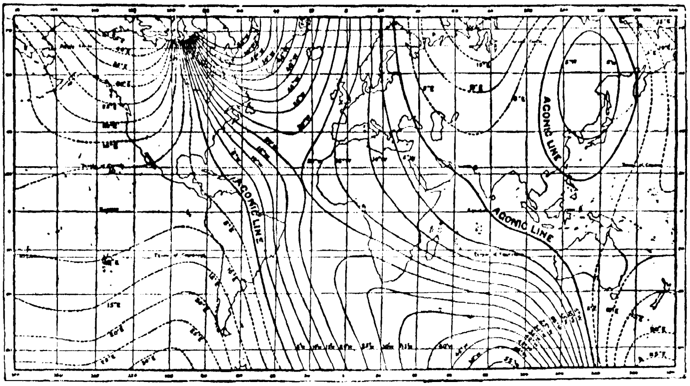


George Philip & Sons, Ltd. FIG. 39. London Geographical Institute

A closed curve formed by the Agonic line in Eastern Asia is called the Siberian Oval.

Fig. 38 shows how the strength of H varies over the earth's surface.

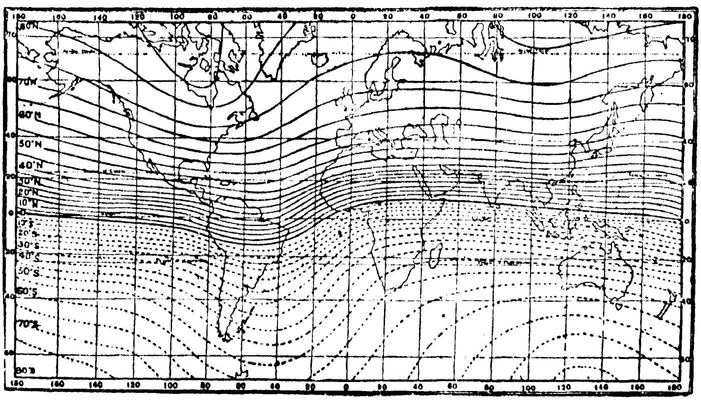
Fig. 39 shows the direction of H over the earth's surface.



George Philip & Sons, Ltd. FIG. 40. *London Geographical Institute*

At any point on one of the lines shown the Magnetic Meridian would be tangential to the line, and the angle between the line and true north would be the declination.

Fig. 40 shows the distribution of the isogonals and agonic lines.



George Philip & Sons, Ltd. FIG. 41 *London Geographical Institute*

Fig. 41 shows the distribution of the isoclinics, and also the aclinic line or magnetic equator.

The Secular Magnetic Variation

If at any definite place the magnetic elements are observed regularly for a long period of time (as they have been at Greenwich), they are found to undergo changes of several kinds. First, and most important, there is a long-period change due apparently to the shifting of the earth's magnetic axis. The declination and dip have changed in a way which may mean that the earth's axis is changing with a period of about five hundred years, or that it is undergoing an irregular change. As observations have not yet been systematically kept long enough for even one complete period, we cannot tell which is true. This variation is known as the secular variation.

The rate of change of the declination is large (so that it has to be allowed for on ships), and does not remain constant. Forty years ago the declination was decreasing at the rate of 5′ per annum, but the rate has increased until it is now about 12′ per annum, and it is at present showing signs of decreasing again. Greenwich Observatory has now ceased to take magnetic observations owing to the introduction of stray fields by the electric trains, and the observation station has been shifted to Abinger, near Dorking, at long. 0° 23′ W., lat. 51° 11′ N.

Here the average values of the magnetic elements for 1927 were :

$$\begin{aligned} \text{Declination} &= 12^\circ\ 58{\cdot}4'\ \text{W.} \\ \text{H} &= 0{\cdot}1858\ \text{oersted.} \\ \text{Dip} &= 66^\circ\ 36'. \end{aligned}$$

The rate of change of H and of the dip are at present negligible.

Periodic Magnetic Variations

The magnetic elements undergo several other variations—a long-period variation depending on the period of the sun-spot cycle (sun-spots reach a maximum about every eleven years); a period of 28 days (or a lunar month); a period of

$26\frac{1}{3}$ days (the period of revolution of the sun's inner core, whose outer surface is not rigid, but revolves at different speeds in different latitudes, all of them being different from that of the inner core); an annual period; and a daily period. Also there are accidental sudden variations called "magnetic storms," usually associated with thunderstorms or intense sun-spots.

Measurement of H

H can be measured by several methods. The most elementary (but the clumsiest) has already been described in detail in the last chapter. Any magnet is used. $\frac{M}{H}$ is determined by any one (or as many as possible) of a number of methods, such as by deflexion magnetometer, by plotting neutral points, or by a vibration magnetometer. MH is determined by finding the period of oscillation of the magnet when freely suspended. H is then the square root of the quotient of MH by $\frac{M}{H}$.

H may also be determined quickly by observing the throw produced in a calibrated ballistic galvanometer when an earth-inductor coil is turned through 180° from a vertical position. This method is described in Chap. V, Part II.

H may also be quickly determined by means of a Tangent galvanometer, or better a Helmholtz galvanometer. The construction and theory of these instruments is explained later, and it is enough now to say that the field at the centre of the Tangent galvanometer is

$$F = \frac{2\pi ni}{10r} \text{ oersted}$$

where n is the number of turns of radius r cm. carrying a current of i amperes.

For the Helmholtz galvanometer, which has a more uniform central field, the value is

$$F = \frac{32\pi ni}{5\sqrt{5}r}.$$

*D

If a deflection of θ is produced in either of these galvanometers, then

$$F = H \tan \theta.$$

So we have for the Tangent galvanometer

$$H = \frac{2\pi ni}{10r \tan \theta} \text{ oersted}$$

and for the Helmholtz galvanometer

$$H = \frac{32\pi ni}{5\sqrt{5}r \tan \theta} \text{ oersted.}$$

Measurement of Declination

It is really a surveyor's job to measure declination. All that is required, of course, is to determine as accurately as possible two directions, and measure the angle between them.

In a well-surveyed country the surveyor first resects (determines from known survey points) his own position accurately, and then uses his known position, and that of whatever survey points he can see, to determine the direction of true north.

In unsurveyed country, when it is impossible to link up with any existing survey system, the surveyor determines true north by stellar or solar observations, with the help of a chronometer or wireless time-signal.

The amateur in a well-surveyed country can get a good idea of true north by orienting a large-scale map carefully, or by observing the direction of the Pole star or of the sun at noon.

The direction of the magnetic meridian can be determined roughly with a prismatic compass, and rather more accurately with a suspended compass-needle armed with sights. It must be remembered that it is impossible to guarantee that the line of the sights is in the same direction as the magnetic axis, since this depends on the accident of the way in which the needle happens to have been magnetized. For this reason the line of the sights must be taken (and marked out by posts in the ground separated as far as possible), both with the needle suspended the right way up and with it upside down. The mean of the two directions observed should be taken. Probably a really good prismatic compass gives the best value in practice.

Measurement of Dip

Dip may be measured directly with an instrument called a Dip-Circle, or indirectly by determining the relative magnitudes of H and either I or V, and calculating θ from $\tan\theta = \frac{V}{H}$ or $\sec\theta = \frac{I}{H}$. The second method, which is the simplest and most artistic, though the less accurate, will be considered first.

If a vibration magnetometer oscillates in time T_H in a horizontal plane where no fields but H affect it, and in time T_I in the plane of the magnetic meridian, then

$$\frac{I}{H} = \left(\frac{T_I}{T_H}\right)^2.$$

The beauty of this method is that only the magnetic meridian need be known, and the direction of I is not known until the result is obtained; for the needle oscillates about the direction of greatest intensity, which *is* the direction of I, automatically.

The time of oscillation due to V may be found by making the needle oscillate in a vertical plane perpendicular to the magnetic meridian.

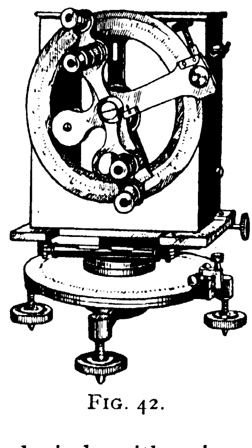


FIG. 42.

The Dip-Circle (Fig. 42) measures θ directly. It consists of a long magnet able to swing quite freely about a horizontal axis perpendicular to its magnetic axis and passing as nearly as possible through its centre of gravity. The axis-spindle is usually supported on agate edges or on nearly frictionless pivots. A vertical graduated circle with micrometer eye-pieces allows the angular deflection from the horizontal of both ends of the needle to be observed accurately. The whole apparatus can turn about a vertical axis, and the direction of the plane in which the needle can turn can be read on a horizontal graduated circle.

The process of finding the dip really consists of two distinct

parts; finding the magnetic meridian and observing the dip-angle. The instrument is turned so that the top end of the needle reads 90° (the needle being approximately vertical). The reading of the horizontal scale is observed.

Usually it will be found that the reading of the bottom end of the needle is now not exactly 90°. The instrument is turned till it is 90° and the horizontal scale reading is observed. The instrument is now turned through 180°, and two similar observations are made. The instrument is then set so that the horizontal scale reading is 90° away from the mean of these four observations; and it should be accurately in the magnetic meridian.

Sixteen readings of the vertical scale are then obtained to give the dip, in the following manner:

Both ends of the needle are read in the first position.

The instrument is turned horizontally through 180° and both ends are read again.

The needle is reversed in its bearings and 4 similar readings are taken, making 8 in all up to now.

The needle is remagnetized so that its poles are changed, and 8 similar readings are taken.

The mean of these 16 readings is the dip.

The intention of all these readings is to counteract all the following sources of error:

The axis of the magnet may not coincide with the centre of the scale.

The axis of rotation may not pass through the centre of gravity.

The magnetic axis of the needle may not coincide with its geometric axis.

Careful consideration of the method will show how these difficulties are overcome by it.

Causes of Terrestrial Magnetism

The three possible causes for the earth's magnetism commonly suggested are:

1. Permanent magnetism in the earth itself.
2. Currents flowing inside the earth.
3. Currents flowing in or above the earth's upper atmosphere.

We will consider these three in order, and show that none is satisfactory in its simple form.

Possibility of Permanent Magnetism

The simplest suggestion is that Fig. 37 roughly represents the truth, and this is supported by the general belief that the earth's core is largely iron. Unfortunately this belief is accompanied by a belief that the temperature of the earth's interior is far above the critical temperature for iron. In fact the temperature at a depth of 20 kilometres probably reaches the critical temperature for iron; and at 15 kilometres it reaches the temperature for magnetite, which has a lower critical temperature than iron.

It therefore follows that either—

(*a*) Any field due to permanent magnetism must be set up by a magnetized shell at the earth's surface not more than 20 km. thick, or

(*b*) The critical temperature of iron must rise with increase of pressure.

The objection to (*a*) is that the intensity of magnetization of this shell would have to be about half the maximum intensity possible in order to produce the existing effect, and no method of accounting for such magnetization has been thought of.

The objection to (*b*) is that such experiments as have been conducted seem to indicate that the critical temperature is, if anything, lowered by rise of pressure. But the experiments are not conclusive.

A serious objection to the hypothesis of permanent magnetism in general is that the sun has a very strong magnetic field although its temperature, even at the surface, is about 6000° C., and it can hardly be supposed to possess permanent magnetism.[1]

Possibility of Earth-currents

Examination of the lines of force round a single coil of wire shows that a current flowing inside the earth under the equator from East to West would produce approximately the right shape of field.

[1] W H Ramsey has suggested that since non-metals become conductors under sufficient pressure (P W Bridgman, Harvard) the cores of the earth and the planets may be conducting even if non-metallic.

But it can be shown that, if at any moment such a current existed, its strength would be reduced to about one-third of its original value in about three days owing to the earth's electrical resistance, which is roughly known. So if such a current existed an expenditure of energy would be required to maintain it, and we do not see where this energy would come from or how it would operate. No simple theory of earth-currents is thus acceptable.

Possibility of Currents in or above the Upper Atmosphere

The reflection of wireless signals with time-intervals extending up to many seconds indicates that the earth is surrounded by conducting layers extending even further out than the moon's orbit. A current flowing from West to East in these layers would (by analogy with the field inside a solenoid) produce a field from South to North in the earth's neighbourhood.

Alternatively an electron-stream from East to West would do this, or a stationary ring or blanket of negative charge with the earth revolving (as it does) from West to East inside it would produce this field with respect to the earth. So superficially our conducting layer (usually called the Heaviside layer, after the great mathematician Oliver Heaviside) looks promising.

Unfortunately, however, it is easy to see that the earth's magnetic field is quite the wrong shape to be mainly accounted for in this way. It curves inwards on itself rather than outwards round the current. So we cannot regard these upper currents (or charges being left behind) as having more than a slight influence on the earth's field.

It has been estimated that less than 1 per cent. of the earth's field can be due directly to upper atmosphere currents, and not more than 2 per cent. to the secondary effect of magnetism induced in the earth by these currents.

The Sun's Magnetic Field

Incandescent gases emit, and cool gases absorb, light of particular frequencies. When observed through a spectroscope, therefore, incandescent gases show bright lines, and cool gases when illuminated from behind by white light show dark lines.

The effect of a magnetic field on the atoms of the gases is to make each frequency produce two or more components close to one another instead of a single one. Each line in the spectroscope thus splits up into components, whose separation enables us to deduce the strength and direction of the magnetic field causing them This effect is known as the Zeeman Effect. A corresponding effect by a strong electric field is known as the Stark Effect. Both names are those of the discoverers.

The Zeeman Effect tells us, through the spectroscope, much about the sun's magnetic field.

It is on the whole like the earth's, being related in the same way as regards general direction to the sun's direction of rotation.

It has, however, very marked differences.

The total intensity of the earth's field is nowhere greater than about 0·65 oersted, or less than about 0·25 oersted. The general intensity of the sun's field reaches values of at least 50 oersted, and in sunspots temporary values of 3000 oersted are not uncommon.

The sun rotates on its axis in about 27·3 days (though it does not go round at the same speed at all latitudes, and thus does not behave as a solid body), but the magnetic axis rotates in 31·4 days. These are the values as apparent to the earth. From the point of view of the sun they would be 26·3 days and 30·4 days. The magnetic axis (which, like the earth's, is tilted with respect to the sun's axis) thus rotates with respect to the sun's axis, in a direction opposite to the sun's direction of rotation, once in about 8 months, whereas it seems likely that the earth's magnetic axis rotates once in about 480 years.

The sun's magnetic intensity can be observed at different levels by observing the Zeeman Effect on different gases, whose levels are approximately known from independent evidence. The lowest measured level is 250 kilometres, at which the intensity is about 50 oersted, but it falls to 10 oersted at a height of 450 kilometres The Zeeman Effect is too small to be observed for fields less than 10 oersted. This sudden falling off is very surprising, and is not, as far as we know, found on the earth. There must be some special reason for it.

Sunspots usually occur in pairs of different magnetic polarity, the line joining them being approximately parallel to the sun's equator. The enormous fields above them do not diminish vertically at anything like the speed of the sun's general field. The polarity of the leading spots of a set of pairs rotating with the sun's surface is usually the same with pairs in the same hemisphere during one sunspot-cycle. It is opposite in the

other hemisphere, and is reversed at the next sunspot minimum.

Further Suggestions of Causes

The closeness of the magnetic axes of both the earth and the sun to their axes of rotation, the general resemblances of the shape of their fields (though there are most important differences as well), and the fact that the direction of the fields are related in the same way to the direction of rotation, suggest that there is at least a close relation between the rotation and the field.

It has been suggested that any large rotating body has a magnetic field because it is rotating, though such an effect is not observed for small rotating bodies. [It is perhaps worth noticing that the ratio of the sun's diameter to the earth's diameter is about the same as the ratio of the maximum intensity of the sun's ordinary field to the maximum intensity of the earth's—about 100 : 1.]

The objection to this is that it can be shown that the external field of a uniformly magnetized sphere is proportional only to the intensity of magnetization, and is independent of the size.

It has also been suggested that there is a cosmic magnetic field, and that currents in the earth and the sun are produced by electromagnetic induction by their rotation in this field; but it is difficult to see why such a field should exist, except to explain terrestrial and solar magnetism!

It has been suggested that in ferromagnetics the actual magnetic elements have angular momentum, and may therefore align themselves like gyroscopes if the body containing them is rotated. The body would then become magnetized along its axis of rotation. Such an effect has actually been found experimentally, but both theoretically and experimentally the resulting field is proportional in intensity to the angular velocity and independent of the size of the sphere. With angular velocities like those of the sun and the earth the effect would apparently be far too small. But the fact that it is proportional to the angular velocity suggests that it is friction which prevents the magnetic elements from turning round, so that the effect might be greater with bodies of high temperature like the earth and the sun, in which the elements are more or less shaken loose by thermal vibrations. But as far as we know the high temperature robs the elements of their magnetism, and thus makes it not worth their while to turn round.

It has been suggested that the earth and the sun are electrically charged bodies, and that the rotation of the charges produces a

magnetic field. It can, however, be shown that enough charge to produce the right strength of field would require a much bigger electric potential gradient at the surface than can be found to exist; and also the field would not have the shape actually observed.

In order to modify this theory to account for the smallness of the electric field near the surface, it has been suggested that the earth has a negative charge on its surface, and an equal positive charge below the surface. But this would produce far too great an electric field inside the earth.

This difficulty could be overcome by supposing an electric polarization of the earth, as distinct from a separation of free charges (*i.e.* a behaviour like an insulator in an electric field rather than like a conductor). This would have the advantage of accounting for the very rapid decrease of the sun's field outwards from the surface, and the great radial electric intensity would help to account for the polarization, but neither of these arguments would apply for the earth.

Lunar Magnetic Variations

Some variations in the earth's field have the same time-intervals as the tides, and are therefore considered to be due to the moon. These variations thus occur nearly twice a day, or 54 times in 28 days. They are considered to be due to tides in the atmosphere like the lunar tides; and variations of barometric pressure having the same frequency can be observed. When such tides occur, the moving body of air has electromotive forces induced in it by its passage across the earth's magnetic field, and the resulting atmospheric currents produce their own magnetic fields in consequence. Presumably similar small effects are produced by tidal motion of the sea across the earth's field, and the induced currents in the sea produce their own fields.

The sun has two separate effects on the lunar variations. It adds its own tidal effects to those of the moon (as it does with the tides of the sea), and it makes the air on the sunlit side of the earth more conducting, because among its rays are some which cause ionization. This makes the induced currents larger, and thus increases the magnetic effects, on the sunlit hemisphere.

The directly tidal effect should perhaps not be regarded as lunar, but should be included under the solar effects.

The sun's effect on the lunar variations is, of course, exactly diurnal, whereas the moon's is not quite diurnal, so the two get out of phase with one another periodically.

Solar Magnetic Variations due to Radiation

These are of several kinds.

The sun, in addition to its tidal effects, produces thermal currents in the atmosphere, which, moving in the earth's field, cause electric currents to be induced and additional magnetic fields to be set up. This effect is about eleven times as great as the lunar effect, and this is how we know it is not tidal; for the sun's tidal effect is less than half that of the moon. These effects thus depend in general on the rate at which the sun is supplying heat. This varies with the seasons, and thus has an annual periodic variation.

Magnetic Variations due to Sunspots

There is evidence that sunspots shoot out streams of charged particles with very high velocity for very great distances, so that the sun behaves rather like those slowly revolving sprinklers (with holes near the axis, not with long arms) which are used for watering lawns. If one looked along one of the streams of particles from the sun, it would appear to curve backwards, because the sun's rotation would leave the particles behind after they were shot out.

When the earth encounters one of these streams of particles the stream probably behaves like a conducting layer travelling downwards across the earth's field. Currents are induced in this layer by their motion across the lines of force. It is probable that the magnetic effect is not due directly to moving charges regarded as currents, because this would imply that the charges were all of the same sign, and, if they were electrostatic, repulsion would probably have scattered them. It is supposed that the stream is on the whole neutral but ionized, and that it has taken a time of the order of 2 days to reach the earth.

The problem of what happens to a charged particle approaching a magnetized sphere is rather difficult. A force, of course, acts on it perpendicular to the direction of its motion and to the direction of the magnetic intensity. It is easy to see that one cannot lightly work out what happens to such a particle approaching the earth's rather complicated magnetic field. The conclusion seems to be that the charged particles should partly separate out, so that some of them go round the earth and come down on the far side. This accounts for the approximate uniformity of some magnetic storms over the earth's surface.

The streams of ions move past the earth at a great pace. The sun swings them right round once in $26\frac{1}{3}$ days with respect to

itself, or $27\frac{1}{3}$ with respect to the earth. The front of the stream thus moves past the earth at

$$\frac{93,000,000 \times 2\pi}{27\frac{1}{3} \times 24 \times 60 \times 60} \text{ miles per second}$$

or 248 miles per second.

The diameter of the earth is 7927 miles, so that it would be plunged completely into the stream in $\frac{7927}{248}$, or 32″.

Magnetic storms do, in fact, spring up all over the earth in a very short time. This fact agrees with the theory outlined, and also suggests that the particles begin to stream out from the sun as the result of some very sudden event, so that the stream of particles has a very sharp edge. Obviously two kinds of periodic variation are associated with this effect.

A storm due to any one spot is likely to recur after 27·3 or 54·6 or 81·9 days, and it is on the average found that storms are more frequently separated by these intervals than by others. Usually of course several spots are active. If there is a magnetic storm it is not probable that the next one will occur after one of these intervals, but it *is* probable that after these intervals a storm will occur.

Secondly, the average intensity of storms will obviously vary with the average intensity of sunspots. Sunspot-cycles are about 11 years long, and the average intensity of storms is found to vary with this period.

Other Variations

The cause of the " secular " variation is not known. Observations have only been taken for less than 100 years, but they indicate a main period of 480 years, with probably another of much longer period superposed. The magnetic poles are probably at their maximum distance from the geographical poles just now.

Thunderstorms, of course, cause intense local variations of short duration, and there are permanent local irregularities in some places owing to the presence of underlying magnetic rock.

The Kelvin Compass

The Kelvin, or Mariner's, compass, consists of a 10″ diameter card or aluminium sheet on which the points of the compass are engraved (see Fig. 43).

The magnets are arranged so as to make the ratio of moment of inertia to magnetic moment as small as possible. A number of light strong magnets fixed together on the axis usually forms the magnetic system. Since $T = 2\pi\sqrt{\frac{I}{MH}}$, the smaller $\frac{I}{M}$ the smaller T, and the quicker the compass is to set in its final position. The card-system is often immersed in methylated spirit, which has the double advantage of reducing the weight on the pivot and increasing the damping.

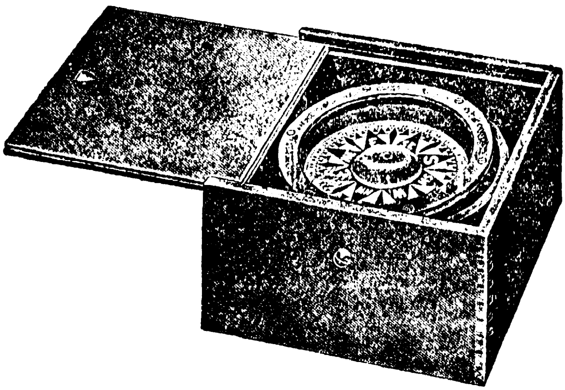


FIG. 43.

Ships' Errors and their Correction. The Quadrantal Error

There are three main kinds of compass-errors on a ship, usually known respectively as the Quadrantal Deviation, the Semicircular Deviation, and the Heeling Error.

The Quadrantal Deviation is due to the magnetization by the earth's field of horizontal pieces of iron on the ship. The direction and magnitude of the error due to such a piece of iron varies with the direction in which the ship is pointing. The deviation is proportional to $\cos\theta_1$, where θ_1 is the angle between the magnetic meridian and the axis of the horizontal bar, and to $\sin\theta_2$ where θ_2 is the angle between the direction of the compass needle and the axis of the bar. As the compass needle very nearly points along the meridian, θ_1 and θ_2 are

almost equal, and the error is thus proportional to sin θ cos θ, or to sin 2 θ Sin 2 θ changes sign four times as θ goes from 0° to 360°, so that the error goes through a half-cycle, from zero to zero, as the ship turns through 90° or one quadrant. This is why this error is called "quadrantal."

It is corrected by placing two soft iron spheres, one each side of the binnacle containing the compass, on the same level as the binnacle. The dimensions and position of the spheres for given errors may be found from Admiralty tables.

The Semicircular Error

The Semicircular Error is due to two effects combined. One is the permanent magnetization of the steel in the ship. This was set up when she was built, and depends on the position she occupied with respect to the meridian while she was being built. The other is the magnetization of vertical pieces of soft iron due to the earth's field.

The error due to the permanent magnetization is proportional to sin θ, where θ is the angle between the magnetic axis of the ship and the compass needle. The error due to vertical magnetization is proportional to tan ϕ where ϕ is the angle of dip.

The combined error due to permanent and vertical magnetization thus varies with (A sin θ + B tan ϕ) where A and B are constants, and thus it goes through a complete half-cycle as the ship turns through 180°. It is therefore called the Semicircular Error.

The permanent magnetization responsible for this error is compensated by small magnets in the compass-box. After the quadrantal error has been corrected, the ship is placed N and S, and small magnets are fixed transversely in the binnacle till the error is compensated. Then the ship is placed E. and W., and more small magnets are put in at right angles to the first lot until the error is again corrected. The correction for vertical magnetization is made by fixing a soft iron bar vertically in front of or behind the binnacle.

The Heeling Error

The Heeling Error is due to the vertical component of the permanent magnetization of the ship. This does not affect

the compass when the ship is on an even keel, but as soon as she heels it does affect it. The effect is greatest when the ship points N. and S., and zero when she points E. and W., unless she not only rolls but pitches.

It is compensated by fixing a permanent magnet near the binnacle, vertical with its axis directly some way below the binnacle, so as to produce at the binnacle a vertical field equal and opposite to the vertical component of the field due to the ship's permanent magnetism.

CHAPTER VIII

THE ELECTRIC CURRENT

The Production of Currents—The Chief Effects of a Current—The Absolute Electromagnetic Units of Current and Charge, the Ampere and the Coulomb—The Absolute Electromagnetic Unit of Potential Difference, and the Volt—Ohm's Law, Resistance, and Conductance—The Absolute Electromagnetic Unit of Resistance, and the Ohm—Conductors in Series—Conductors in Parallel—Specific Resistance and Specific Conductance—Potential Difference and Electromotive Force—Kirchoff's Laws for Steady Current Circuits—Cells in Series and Parallel—Miscellaneous Problems.

The Production of Currents

It was found by Volta in 1799 that if two dissimilar metals (such as copper and zinc) were separated by a wet cloth in contact with both of them, a potential difference existed between them; and that if they were connected by a conductor this potential difference was, to a certain extent, maintained. In this way was obtained for the first time a conductor which was not at the same potential all over, but which had an electric intensity inside it, causing a continuous flow of charge. In order to keep this current flowing, energy had to be supplied continuously; and this energy came from chemical decomposition of the metallic elements. [Cells will be considered in detail later. They are only introduced here because they were the origin of the tremendous discovery of the electric current; a discovery which not only linked the hitherto separate Theories of Static Electricity and Magnetism, but also revolutionized the world in which we live.]

The Chief Effects of a Current

Currents were found to have three major effects, magnetic, thermal, and chemical. Firstly, when flowing they produced a magnetic field, which proved to be related to them in the oddest manner: a manner indeed so odd that it was only

discovered (accidentally while playing about with the apparatus on his bench with some students at the end of a lecture) by Oersted in 1820, twenty-one years after Volta produced a current. The magnetic field acts cylindrically round the wire carrying the current, in the manner of Fig. 44. The vertical arrow gives the direction of the current, flowing from positive to negative, and the arrows in the circles show the direction of the field, being the direction in which it would exert force on a north pole.

The current and magnetic field are related by various rules. The Right-Hand rule states that if the wire carrying the current is grasped with the right hand so that the thumb lies

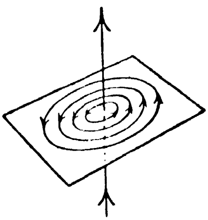


FIG. 44.

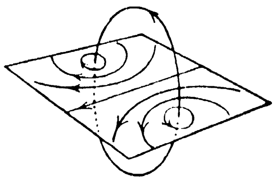


FIG. 45.

along the wire and the fingers encircle it, then if the thumb points in the direction of flow of the current from positive to negative, the fingers give the direction of the magnetic field encircling the current.

The name Right-Hand rule is unfortunate, because there also exists Fleming's Right-Hand rule for the direction of the E.M.F. induced by the motion of a conductor across a magnetic field.

Alternatively, if an ordinary corkscrew be supposed to be penetrating a cork in the direction of the current, a point on the handle moves (if one neglects the forward motion into the cork) in the direction of the magnetic field (but there exists a Scotsman with a left-handed corkscrew which vitiates this rule). Nevertheless, to avoid confusion, the rule will be known as the Corkscrew Rule.

Fig. 45 illustrates the cross-section of the magnetic field due to the combined effect of the two limbs of a circular current.

The main thing to notice about this magnetic field is that at the centre of the circle it acts perpendicularly to the plane of the circle. This fact is made the basis of the definition of the electromagnetic unit of current.

The magnetic effect of currents, and the forces associated with it, will be considered in detail in the chapter on Electromagnetism and Electromagnetic Induction; for the present the effect will only be used to give the definition of the Absolute Electromagnetic Unit of Current, and the Ampere.

Secondly, currents are found to produce heat. The relation between the heating effect and the magnetic effect is surprising at first when it is realized. For it is found that if two currents A and B are so related that B produces just double the magnetic field of A at any given external point, then B produces heat at not double but four times the rate of A. Are we then (when attacking the problem from first principles) to regard B as twice A, or as four times A, or as some other multiple? This problem is discussed when the units of current are defined.

Thirdly, a current flowing through the solution of a metallic salt (called an electrolyte) deposits the metal on the terminal plate at which it leaves the electrolyte, at a definite rate. This rate of deposition is found to be proportional to the strength of the magnetic field due to the current at any given external point. At this point, then, we may say that, if we are choosing between the magnetic field and the heating effect to define our unit current, it is at the moment two to one on the field.

The second effect of a current is discussed in detail in the chapter on Heating Effects, and the third in the chapter on Electrolysis.

The Absolute Electromagnetic Units of Current and Charge, the Ampere and the Coulomb

The justification for the choice of the magnetic, rather than the heating, effect for measuring current is really Rowland's experiment, although the choice was made by Ampère in 1823, and H. A. Rowland published his results in 1889.

It is important to realize that we must find, not which effect is more convenient for measurement, but which effect

(if either) gives a value which, measured electromagnetically, is proportional to the rate of flow of electrostatic charge.

Rowland rotated an ebonite disc, having alternate sectors charged and uncharged. The charged sectors were gilt and the uncharged bare. The uncharged sectors acted as barriers to keep the charge fixed with respect to the disc.

These moving electrostatic charges formed currents which were proportional to the rate at which they were moving. They set up a magnetic field which was measured with a magnetometer placed below the disc. It was found that this magnetic field was proportional in intensity to the rate of revolution of the disc, and thus to the current. Ampère's choice sixty-six years before was thus justified.

Ampère's theorem, which deals with the measurement of current by its magnetic effect, is rather mathematical, and is given in some detail in Part II. Luckily a simple deduction from it may be used to define the absolute unit. The definition which follows is therefore not the right one, but it is the best we can have without calculus.

"The Absolute Electromagnetic Unit of Current is that current which, flowing in 1 cm. of wire bent into the arc of a circle of 1 cm. radius, produces a field of 1 oersted at the centre of the circle." This unit is sometimes called the Abamp.

It is obvious that the effect at the centre is cumulative, being due to every bit of the wire, and thus to the length. So 1 cm. of wire must be chosen. It also depends on the distance away of every bit, so that every bit must be the same distance away, 1 cm., in the same direction. So an arc of a circle of 1 cm. radius is necessary.

This unit of current is found in practice to be inconveniently large; so the ampere, or practical unit, is chosen to be one-tenth of the absolute unit.

The absolute Electromagnetic Unit of Charge is, of course, the quantity of charge which passes when the unit of current flows for one second.

The ratio of the electromagnetic units of current and charge to the electrostatic units is found by measurement to be 3×10^{10}.

The practical unit of charge is the coulomb, which is $\frac{1}{10}$ of the absolute unit, and thus also 3×10^9 electrostatic units.

The Absolute Electromagnetic Unit of Potential Difference, and the Volt

It is a simple matter to show that the potential difference between Volta's metallic elements is of the same nature as the potential difference between the plates of a charged condenser, although for a small pair of elements it is difficult to measure with an electroscope. There are, however, three easy ways of doing it.

Firstly we can take a great number of voltaic elements (forming thus a voltaic pile), earth one end of the pile and connect the other to the knob of an electroscope. If we have enough elements we shall get a deflection.

Secondly we can charge a condensing electroscope with a single element, and obtain a deflection by reducing the capacity of the electroscope.

If we then move the plate away from earth we reduce the capacity without altering the charge: thus increasing the potential and causing the leaves to be deflected.

Thirdly, as the potential difference between the Volta plates is of the order of 1 volt, it can be measured directly with the quadrant (or better the Lindemann) electrometer. This is the best way. The electromagnetic unit of potential is defined as follows :—

"When 1 erg of work is done in taking 1 electromagnetic unit of charge from a point A to a point B, the potential difference between A and B is 1 electromagnetic unit."

Similarly the practical unit, the volt, is defined as follows :—

"When 1 joule of work is done in taking 1 coulomb from A to B, the potential difference between A and B is 1 volt."

One may also express this by saying that the voltage between A and B is equal to the number of joules of work done between A and B per ampere per second.

A joule is simply the practical unit of work or energy. The erg is inconveniently small, and the joule is chosen as 10^7 ergs. We can show that the volt is 10^8 absolute electromagnetic units of potential by a simple argument.

An ampere, which is $\frac{1}{10}$ of the absolute unit, does 10^7 ergs of work per second in flowing across 1 volt, so the absolute unit does 10^8 ergs per second in flowing across 1 volt; so 1 volt is 10^8 absolute units of potential.

Ohm's Law, Resistance, and Conductance

Ohm's Law states that—

"The steady current flowing through a metallic conductor at constant temperature is proportional to the potential difference between the ends of the conductor."

This law, the most familiar and useful of all electrical laws in the affairs of daily life, apparently so obvious that some people are almost surprised to see it stated as a law, has the most amazing history. Georg Simon Ohm published it in 1827, after several years' research, and it was received with

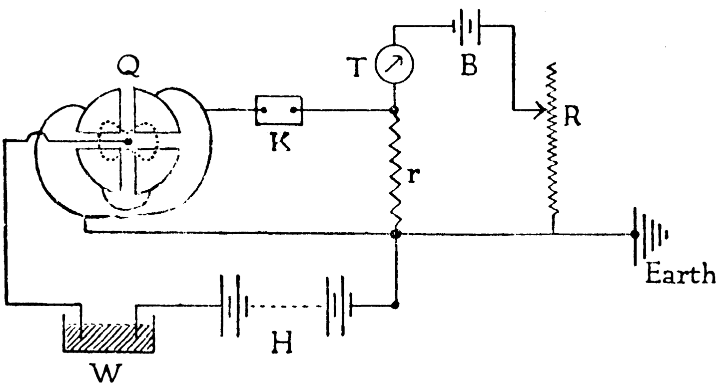


Fig. 46.

scorn and derision. It was described as "a web of naked fancies," and Ohm was told by the Minister of Education that "a physicist who professed such heresies was unworthy to teach science." He did, in fact, lose his teaching post. The law was not generally accepted till 1833, and he received the Royal Society medal for it in 1841.

It may be directly verified from first principles with a Tangent Galvanometer (described in the next chapter), which uses the definition of the absolute unit to measure current, and the Quadrant Electrometer.[1] The circuit is arranged as in Fig. 46, in which Q is a quadrant electrometer, T a tangent galvanometer, *r* a fixed resistance, R a variable resistance, W a water-resistance, H a high-tension battery, K a key, B a small battery of two or three cells.

[1] Besides carrying some current robbed from *r*, an ordinary voltmeter in its construction assumes Ohm's Law which the experiment is designed to demonstrate.

When the key K is in, the deflection d of the electrometer is proportional to the difference of potential between the ends of r. The current through r passes through T also, and is proportional to $\tan\theta$ if θ is the deflection of T. It is found experimentally that $\tan\theta$ is very closely proportional to d, so long as the earth's magnetic field, which determines the deflection of T, is constant.

The meaning of the term Resistance depends upon the truth of Ohm's Law.

Resistance is defined as the ratio of the potential difference across a conductor to the current through it. It is a constant for a metallic conductor carrying steady currents at constant temperature.

Conductance is the reciprocal of resistance: the ratio of the current flowing through a conductor to the potential difference across its ends.

It is important to realize (and apt to be forgotten) that logically resistance appears after the units of current and potential difference have been settled; and that it never would have been conceived if Ohm's Law had not been formulated and accepted.

The Absolute Electromagnetic Unit of Resistance, and the Ohm

"The Absolute Electromagnetic Unit of Resistance is the resistance of a conductor when 1 absolute unit of potential difference across it causes 1 unit of current to flow through it."

"The Ohm is the resistance of a conductor which passes 1 ampere when the potential difference across it is 1 volt."

It follows from these definitions that 1 ohm $= 10^9$ absolute units; for (by its definition) it passes $\frac{1}{10}$ unit of current for a potential difference of 10^8 units; so that it passes $\frac{1}{10^9}$ unit of current for a potential difference of 1 unit.

For legal purposes the ohm is defined as the resistance of a column of mercury 106·300 cm. long, of mass 14·4521 grammes, of constant cross-section, at 0° C., to an unvarying current. The unvarying current is measured electrolytically, as shown in Chapter XI. Two separate legal standards are, of course, required. Those chosen are the ohm and the ampere.

By Ohm's Law, if a current i flows through a resistance R which has a potential difference V across its ends, then

$$R = \frac{V}{i} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

or, Resistance = Potential Difference over current.

This, the fundamental equation of Ohm's Law, may be written :

$$i = \frac{V}{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

or, Current = Potential Difference over resistance, and again,

$$V = Ri \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

or Potential Difference = Current × Resistance.

This obvious equation is given in all three forms because it is surprising how often it is not realized that there are three forms.

Conductors in Series

If n conductors, of resistance, R_1, R_2, R_3, . . . R_n respectively, are arranged in series, and a current i flows through them, then the potential differences across them are R_1i, R_2i, R_3i, . . . R_ni. So the total potential difference is, by equation (3) of Ohm's Law,

$$R_1i + R_2i + R_3i + \ldots + R_ni.$$

or

$$(R_1 + R_2 + R_3 + \ldots + R_n)i.$$

Thus the total potential difference is that due to a current i flowing through a conductor of resistance R, where

$$R = R_1 + R_2 + R_3 + \ldots + R_n \; . \quad . \quad . \quad . \quad (4)$$

This is then the total resistance of all the original conductors in series.

Thus it follows that the total resistance of any number of conductors arranged in series is equal to the sum of their separate resistances.

Since conductance is the reciprocal of resistance, it follows that if C be the total conductance, and C_1, C_2, C_3, . . . C_n, the separate conductances,

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots + \frac{1}{C_n} \quad . \quad . \quad . \quad (5)$$

Thus the reciprocal of the conductance of a set of conductors in series is the sum of the reciprocals of their conductances taken separately.

Conductors in Parallel

If the same set of conductors were arranged in parallel (Fig. 47) and a current i were sent through them, dividing itself up as it passed into $i_1, i_2, i_3, \ldots i_n$, then

$$i = i_1 + i_2 + i_3 + \ldots + i_n.$$

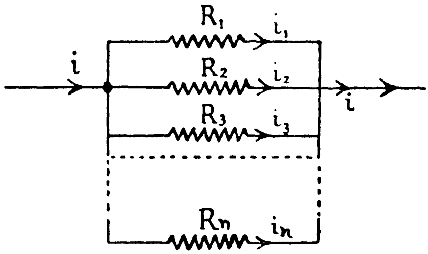


FIG. 47.

If V is the potential difference across the ends of all the paralleled conductors, and R their combined resistance in parallel, then by equation (2) of Ohm's Law

$$\frac{V}{R} = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3} + \ldots + \frac{V}{R_n}$$

So
$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots + \frac{1}{R_n} \quad . \quad . \quad . \quad (6)$$

It thus follows that the reciprocal of the resistance of a number of conductors in parallel is the sum of the reciprocals of their separate resistances.

Similarly, since the conductance C is the reciprocal of the resistance R_1 the conductance C of all the conductors in parallel is given by

$$C = C_1 + C_2 + C_3 + \ldots + C_n \quad . \quad . \quad . \quad (7)$$

So that the conductance of a set of conductors in parallel is the sum of their separate conductances.

Any one of a set of conductors in parallel is said to be a " shunt " for the others, or to " shunt " the others.

Specific Resistance and Specific Conductance

The Specific Resistance of a material is the resistance between the opposite faces of a centimetre cube of the material. It is

so chosen in order to give a quantity depending solely on the nature, and not at all on the dimensions, of the material. Specific resistance satisfies this need, since its calculated value from measurement of any piece of the material is the same, whatever the dimensions of the material.

Let us consider the relation between the resistance and specific resistance of a piece of wire. Let its specific resistance be S, its resistance R, its length L, and its cross-section A.

Imagine first a centimetre cube of the material. Its resistance from face to face is S. Imagine it cut into 100 lengths, each 1 cm. long from face to face, and of cross-section $\frac{1}{100}$ sq. cm. The conductance of each is $\frac{1}{100}$ of the whole. So its resistance is 100 S (or $S/\frac{1}{100}$).

Similarly, the resistance of a 1 cm. length of cross-section A is $\frac{S}{A}$. What now is the resistance of L of these lengths in series? It is L times as much as that of 1 cm. length.

So
$$R = \frac{LS}{A} \quad \quad (8)$$

Similarly
$$S = \frac{RA}{L} \quad \quad (9)$$

Since conductance is the reciprocal of resistance, the specific conductance K is given by

$$\frac{1}{K} = \frac{A}{L} \cdot \frac{1}{C}$$

or
$$K = \frac{CL}{A} \quad \quad (10)$$

It is important to realize that Specific Resistance is NOT the resistance of a cubic centimetre, but that *between opposite faces of a centimetre cube.* A cubic centimetre could take the form of a long thin wire of high resistance, or a disc of negligible resistance.

Potential Difference and Electromotive Force

These quantities, both of which are normally measured in volts, must be clearly distinguished in steady current circuits.

The " Potential Difference " between any two points in a circuit is the algebraic sum of the products of current and

resistance along any path joining the two points in the circuit, provided that no source of power, such as a dynamo or cell, occurs along the path.

The "Electromotive Force" is the algebraic sum of the products of current and resistance along any path going round the whole circuit.

Thus the Potential Difference is localised as occurring between two definite points, whereas the Electromotive Force is not localised, but belongs to the whole circuit.

Alternatively, the Electromotive Force may be regarded as the entity which causes the electric charge to circulate.

In this sense an individual cell may be regarded as having an Electromotive Force which is independent of the rate at which current is taken out of the cell, and which can be determined by allowing the cell to discharge through a resistance, and measuring the algebraic sum of current × resistance round the circuit.

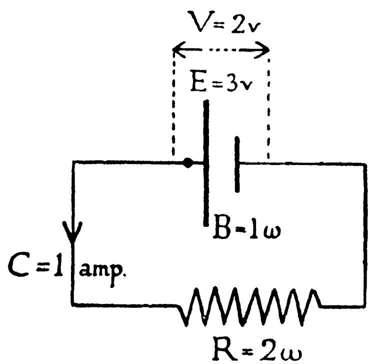


FIG. 48.

If a circuit has several cells in it, the total E.M.F. in the circuit is the algebraic sum of the E.M.F.s of the separate cells. If the Potential Difference is being observed for a part of the circuit which contains one or more sources of power, such as cells or dynamos, then it is equal to the algebraic sum of the products of current and resistance along the path, and the E.M.F.s of any sources of power along the path. The whole problem becomes clearer if we consider a circuit containing a cell having internal resistance and an external resistance.

In Fig. 48, let a cell of internal resistance 1 ohm drive a current of 1 amp. through an external resistance of 2 ohms. All these quantities can be directly measured by methods given in the next chapter.]

By our definition of P.D., the P.D. between the terminals of the cell $= iR = 1 \times 2$ volts $= 2$ volts.

By our definition of E.M.F., the E.M.F. of the circuit (or of the cell, since it is the only source of power in the circuit) is

$$i(B + R) = 1 \times (1 + 2) = 3 \text{ volts.}$$

If the cell is producing current, this 3 volts cannot be observed directly by any instrument, but can only be deduced; but it is none the less a most useful constant of the circuit. The P.D. (of 2 volts in this case) can always be observed directly by any instrument which measures potential.

It is clear from the calculations that

$$i = \frac{V}{R} = \frac{E}{B + R} = \frac{E - V}{B} \quad . \quad . \quad . \quad (11)$$

This very important set of equations can be used to solve all simple problems involving a single cell in a circuit.

It is worth noticing that since

$$\frac{E}{V} = \frac{R + B}{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

$$\text{As } R \longrightarrow \infty, \ \frac{R + B}{R} \longrightarrow 1, \text{ and } V \longrightarrow E.$$

Thus on open circuit, when we may regard R as infinite, the potential difference is equal to the E.M.F. This gives the quickest method of finding the E.M.F.

Kirchoff's Laws for Steady Current Circuits

Kirchoff's Laws are :—

1. At any junction in a circuit the algebraic sum of currents flowing into the junction is zero; or, inflowing current equals outflowing current. [In other words, there is no accumulation of charge at any point in the circuit.]

2. In a closed circuit, the algebraic sum of the products of current and resistance of each part of the circuit, taken along any path, is equal to the algebraic sum of the electromotive forces along that path.

These laws introduce nothing new. The first is applied common sense, and the second an extension of Ohm's Law,

Equation No. 3. The operation of these laws is best understood by working out a number of examples.

Example 1.—A battery of E.M.F. 4 volts sends 2 amps. through a 1-ohm coil. What current would it supply (*a*) if short-circuited, and (*b*) through a 2-ohm coil?

The arrangement of the circuit is as in Fig. 48, but the numbers are different.

We know E, *i*, and R.

By the 2nd law or by equation (11)

$$E = i(B + R)$$
$$4 = 2(B + 1)$$
$$\therefore\ B = 1 \text{ ohm.}$$

(*a*) If $R = 0$, we have

$$4 = i(1 + 0)$$
$$\text{so } i = 4 \text{ amps.}$$

(*b*) If $R = 2$, we have

$$4 = i(1 + 2)$$
$$\text{so } i = 1\tfrac{1}{3} \text{ amps.}$$

Any problem involving this circuit can be easily solved by the equations (11) and (12) at the end of the section on Electromotive Force and Potential Difference. These equations are clearly applications of Kirchoff's 2nd law.

Example 2.—A Daniell cell of E.M.F. 1·1 volts, and resistance 0·4 ohm, is connected in parallel with a Leclanché cell of E.M.F. 1·5 volts and resistance 1 ohm. They are made to send a current through an external resistance of 2 ohms. Find how much this current is, and how much each cell supplies.

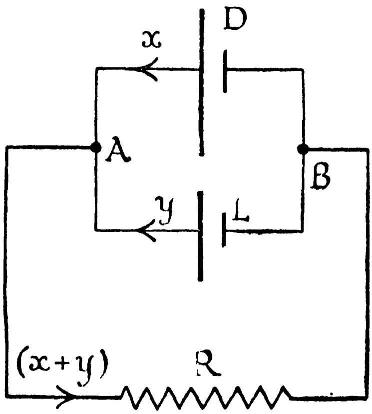


FIG. 49.

Let D (Fig. 49) be the Daniell, L the Leclanché, and A and B the points of junction in the circuit.

Let the current from the Daniell be x amps., and that from the Leclanché y amps.

Then, by Kirchoff's first law, the current flowing out of the junction through R, the external resistance, is $(x + y)$ amps.

Applying Kirchoff's second law for the complete circuit BDAR, in which the only E.M.F. is that of the Daniell,

$$0{\cdot}4x + 2(x + y) = 1{\cdot}1.$$

Applying the law to the circuit BLAR, in which the only E.M.F. is that of the Leclanché,

$$1{\cdot}0y + 2(x + y) = 1{\cdot}5.$$

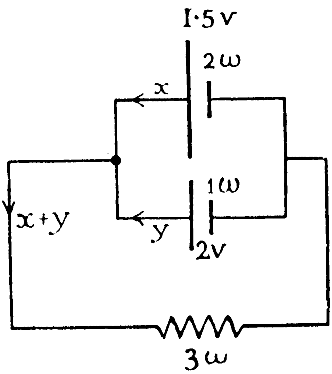


FIG. 50.

[We could also apply the law to the circuit BDAL, in which both the E.M.F.s of Daniell and Leclanché occur, opposing each other. We should get :

$$1{\cdot}0y - 0{\cdot}4x = 1{\cdot}5 - 1{\cdot}1,$$

which is simply the equation obtained by subtracting the BDAR equation from the BLAR equation. So it tells us nothing new.]

Rearranging the first two equations we get :

$$2{\cdot}4x + 2y = 1{\cdot}1$$
$$2x + 3y = 1{\cdot}5.$$

Solving these, we find that x, the current from the Daniell, is $\frac{3}{32}$ amp., and y, the current from the Leclanché, is $\frac{7}{16}$ amp., so that the total current flowing through the external resistance is $\frac{17}{32}$ amp.

Example 3.—Cells of E.M.F.s 1·5 volts and 2 volts, and internal resistances 2 ohms and 1 ohm respectively, send currents x and y through an external resistance of 3 ohms. Find x, y, and the main current. See Fig. 50.

By the 1st law, the external current is $(x + y)$ through the 3-ohm resistance.

Taking the 2nd law equation for the circuit containing the 1·5-volt cell and the external resistance—

$$2x + 3(x + y) = 1{\cdot}5.$$

Taking the 2nd law equation for the circuit containing the 2-volt cell and the external resistance—

$$y + 3(x + y) = 2.$$

If we take the 2nd law equation for the circuit containing both cells, but not the external resistance, we shall find we do not get an independent equation, but simply the equation obtained by subtracting one of our first two equations from the other. Obviously, as we only have two unknowns, x and y, we cannot get more than two independent equations.

Rewriting our two equations—

$$\begin{aligned} 5x + 3y &= 1\cdot5 \\ 3x + 4y &= 2. \end{aligned}$$

The solution of these equations is

$$\left.\begin{aligned} x &= 0 \\ y &= \tfrac{1}{2} \text{ amp.} \end{aligned}\right\}$$

so the 2nd cell supplies all the current.

Let us now alter the internal resistance of the top cell from 2 ohms to 0·5 ohm.

Our original equations become

$$\begin{aligned} 0\cdot5x + 3(x + y) &= 1\cdot5 \\ y + 3(x + y) &= 2. \end{aligned}$$

which reduce to

$$\begin{aligned} 7x + 6y &= 3 \\ 3x + 4y &= 2. \end{aligned}$$

Again we get

$$\left.\begin{aligned} x &= 0 \\ y &= \tfrac{1}{2} \text{ amp.} \end{aligned}\right\}$$

This rather surprising result appears because the potential difference across the more powerful cell, when it is sending the same current as it would send unaided through the external resistance, is exactly equal to the E.M.F. of the less powerful cell. So the latter can never send any current, and it does not therefore matter what its internal resistance is.

Let us now alter the E.M.F. of the first cell from 1·5 volts to 3·0 volts, leaving its resistance 2 ohms as it was originally. The equations become

$$\begin{aligned} 2x + 3(x + y) &= 3 \\ y + 3(x + y) &= 2, \end{aligned}$$

which reduce to

$$5x + 3y = 3$$
$$3x + 4y = 2.$$

The solution of these is

$$\left.\begin{aligned} x &= \frac{6}{11}\text{ amp.} \\ y &= \frac{1}{11}\text{ amp.} \end{aligned}\right\}$$

The main current is thus $\frac{7}{11}$ amp.

Example 4.—The diagram (Fig. 51) shows an unbalanced Wheatstone bridge, through which current is driven by a 2-volt accumulator of negligible resistance.

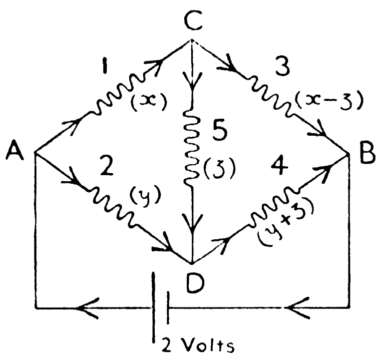


FIG. 51.

Find the currents in all the branches, and the effective resistance of the whole network to the current from the accumulator.

Let x, y, z, be the currents in the branches AC, AD, CD, and their directions those of the arrows.

By Kirchoff's first law, the current in CB is $(x - z)$, and in DB is $(y + z)$, flowing out from C and D respectively. Since the potential drop from A to B is 2 volts we have by the route ACB:

$$1(x) + 3(x - z) = 2;$$

by the route ADB:

$$2(y) + 4(y + z) = 2;$$

by the route ACDB:

$$1(x) + 5(z) + 4(y + z) = 2.$$

If we try the route ADCB we shall merely get an equation which is not independent, but could be obtained from the three already given.

Rearranging these equations we get:

$$4x - 3z = 2$$
$$6y + 4z = 2$$
$$x + 4y + 9z = 2.$$

Solving, we obtain :

$$x = \frac{44}{85} \text{ amp.}$$

$$y = \frac{27}{85} \text{ amp.}$$

$$z = \frac{2}{85} \text{ amp.}$$

which will be found to satisfy the equations.

The total current flowing from the accumulator is thus $\frac{44 + 27}{85}$, or $\frac{71}{85}$ amp.

So the total effective resistance of the network is $\frac{2}{\frac{71}{85}}$, or $2\frac{28}{71}$ ohm.

Example 5.—11 equal wires, each of resistance 5 ohms, are laid along 11 of the edges of a non-conducting cube, their ends being soldered together to form joints of no resistance. Prove that the total resistance from one end of the vacant edge of the cube to the other is 7 ohms.

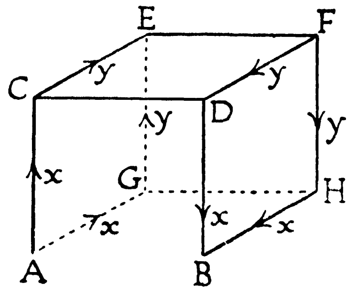


FIG. 52.

Fig. 52 shows the arrangement. We are required to find the resistance to a current flowing from A to B.

By symmetry AC, AG, DB, HB, all carry the same current x. By symmetry CE, GE, FD, FH, all carry the same current y. By Kirchoff's 1st law, the current in CD and GH (the same by symmetry) is $(x - y)$, and the current in EF is $2y$.

Since we only have two unknowns, x and y, we only need two independent equations connecting them. Let the E.M.F. from A to B be E volts, supplied from a source of no internal resistance.

Then, by Kirchoff's 2nd law—
Along the path ACDB

$$E = 5x + 5(x - y) + 5x = 15x - 5y.$$

Along the path ACEFDB

$$\begin{aligned} E &= 5x + 5y + 10y + 5y + 5x \\ &= 10x + 20y. \end{aligned}$$

$$\begin{aligned} \therefore\ 15x - 5y &= 10x + 20y \\ 5x &= 25y \\ y &= \frac{x}{5}. \end{aligned}$$

$\therefore$ from the first equation

$$E = 15x - \frac{5x}{5} = 14x.$$

But the total current flowing is $2x$.

So the total resistance $= \dfrac{E}{2x} = 7$ ohms, since $E = 14x$.

Cells in Series and Parallel

The E.M.F. of a set of cells in series is the sum of their separate E.M.F.s, and their combined resistance is the sum of their separate resistances. These facts are obvious.

The resistance of a set of similar cells in parallel obeys the same rules as any other set of equal resistances in parallel. The resistance of n cells in parallel, each of resistance B, is $\dfrac{B}{n}$.

Arranging n cells in parallel is really just the same as multiplying the area of the plates of one cell by n; so it makes no difference to the E.M.F.

Thus if we have n cells in parallel, each of resistance B and E.M.F. E, then their combined E.M.F. is still E, but their combined resistance is $\dfrac{B}{n}$, whereas, if we put them in series, their combined E.M.F. would be nE, and their combined resistance nB.

These facts bring us to the simple problem of how best to arrange a given number of cells so that they should send the largest possible current through a given resistance. Let us consider first the case of a very small resistance, so small that it may be taken as a short circuit.

1 cell would give a current $\dfrac{E}{B}$ if short-circuited.

n cells would give $\dfrac{nE}{nB}$, or $\dfrac{E}{B}$ again, if connected in series.

n cells would give $\dfrac{E}{\frac{B}{n}}$, or $\dfrac{nE}{B}$, if connected in parallel.

Connection in parallel thus gives the largest current for a very small external resistance.

Let us consider a very large external resistance—so large that the resistance of all the n cells in series is negligible compared with it. Let this resistance be R.

Then the current with the cells in series is very nearly

$$\frac{nE}{R}, \text{ or accurately } \frac{nE}{R + nB}.$$

Whereas with the cells in parallel it is very nearly

$$\frac{E}{R}, \text{ or accurately } \frac{E}{R + \frac{B}{n}}.$$

Consequently we have the general rule that, the bigger the external resistance, the more the cells should be in series to give a big current, and, the smaller the external resistance, the more the cells should be in parallel to give a big current.

It is shown in Appendix 3, Part II, by the methods of calculus, that if n cells of resistance B are required to send a maximum current through an external resistance R, then they should be arranged in $\frac{n}{p}$ series sets of p cells in parallel, where p is one of the two integers nearest to $\sqrt{\frac{nB}{R}}$.

There is always a particular value of the external resistance which gives the same current for cells in series or parallel. This is the value for which the series current $\frac{nE}{R + nB}$ is equal to the parallel current $\frac{E}{R + \frac{B}{n}}$.

If we wish to find it, we have

$$\frac{nE}{R + nB} = \frac{E}{R + \frac{B}{n}}$$

$$\therefore\ nR + B = R + nB$$

$$R = B$$

—a surprisingly simple result.

If the external resistance is more than that of one cell, then the cells give the bigger current when they are arranged in series.

*E

If the external resistance is less than that of one cell, then the cells give the bigger current when they are arranged in parallel. Cells can, of course, be arranged also in series-parallel, which means a series of sets of paralleled cells.

Example 6.—Six Daniell cells have an E.M.F. of 1·1 volts each, and an internal resistance of 0·4 ohm each. Find the current they will send through an external resistance of (*a*) 0·2 ohm, (*b*) 1·0 ohm, when they are arranged in series, in parallel, and in both possible arrangements of series-parallel.

(*a*) External resistance 0·2 ohm.

6 cells in series have E.M.F. 6·6 volts and internal resistance 2·4 ohms.

$$\text{Current} = \frac{6{\cdot}6}{0{\cdot}2 + 2{\cdot}4} = \frac{6{\cdot}6}{2{\cdot}6} = 2{\cdot}54$$

6 cells in parallel have E.M.F. 1·1 volts and internal resistance 0·0667.

$$\text{Current} = \frac{1{\cdot}1}{0{\cdot}2 + 0{\cdot}0667} = \frac{1{\cdot}1}{0{\cdot}2667} = 4{\cdot}12 \text{ amps.}$$

6 cells, 2 series sets of 3 in parallel, have E.M.F. 2·2 volts, and internal resistance $2 \times \frac{0{\cdot}4}{3} = 0{\cdot}2667$ ohm.

$$\text{Current} = \frac{2{\cdot}2}{0{\cdot}2 + 0{\cdot}2667} = \frac{2{\cdot}2}{0{\cdot}4667} = 4{\cdot}70 \text{ amps.}$$

6 cells, 3 series sets of 2 in parallel, have E.M.F. 3·3 volts, and internal resistance $3 \times \frac{0{\cdot}4}{2} = 0{\cdot}6$ ohm.

$$\text{Current} = \frac{3{\cdot}3}{0{\cdot}2 + 0{\cdot}6} = \frac{3{\cdot}3}{0{\cdot}8} = 4{\cdot}125 \text{ amps.}$$

(*b*) It is easy to prove in the same way that the corresponding values, for an external resistance of 1 ohm, are, taken in the same order, 1·94 amps., 1·03 amps., 1·74 amps., 2·06 amps.

The general problem, which involves calculus, is considered in Appendix 3, Part II.

Miscellaneous Problems

Example 7. *Potential Drop.*—(*a*) A power-station supplies 500 amps. at 200 volts, through mains two miles long having a resistance of 0·01 ohm per mile.

What is the potential drop in the mains?

What percentage of the power supply is doing useful work?

There are two mains, for outgoing and incoming current, involving 4 miles length altogether, so that the total resistance is 0·04 ohm.

$$\therefore \text{ Potential Drop} = 500 \times 0{\cdot}04 = 20 \text{ volts,}$$
$$\text{Useful P.D.} = 200 - 20 = 180 \text{ volts,}$$

$\therefore \dfrac{180}{200} \times 100\%$, or 90% of the power supply does useful work.

(*b*) If 300 amps. had been taken half-way along the mains, and 200 amps. at the end, what is the effective P.D. at both points of supply?

Potential Drop down first half $= 500 \times 0{\cdot}02 = 10$ volts,

$$\therefore \text{ P.D. half-way} = 200 - 10 = 190 \text{ volts.}$$

Additional Potential Drop in second half $= 200 \times 0{\cdot}02 = 4$ volts,

$$\therefore \text{ P.D. at the end} = 190 - 4 = 186 \text{ volts.}$$

Example 8.—Resistances in parallel.

Resistances of 8 and 12 ohms are in parallel.

(*a*) If the 12-ohm resistance carries 4 amps., what do both carry together?

(*b*) If a resistance R is joined in parallel with the first two, and the total current is the same though the 12-ohm resistance now only carries 3 amps., what is R?

(*a*) The P.D. across the 12-ohm resistance $= 12 \times 4 = 48$ volts,

$$\therefore \text{ current through 8-ohm resistance} = \frac{48}{8} = 6 \text{ amps.,}$$

$\therefore$ total current $= (6 + 4) = 10$ amps.

(*b*) The P.D. across the 12-ohm resistance $= 12 \times 3 = 36$ volts,

$$\therefore \text{ current through 8-ohm resistance} = \frac{36}{8} = 4\tfrac{1}{2} \text{ amps.,}$$

$$\therefore \text{ current through R} = 10 - 3 - 4\tfrac{1}{2} = 2\tfrac{1}{2} \text{ amps.,}$$

$$\therefore \text{ R} = \frac{36}{2\frac{1}{2}} = \frac{72}{5} = 14{\cdot}4 \text{ ohms.}$$

Example 9.—Inaccurate reading of low-resistance voltmeter. The potential difference between the terminals of a cell is $1\frac{3}{7}$ volts when it is measured with a voltmeter of internal resistance 1000 ohms, and 1 volt when it is measured with a voltmeter of internal resistance 100 ohms.

Find the E.M.F. (E) and internal resistance (B) of the cell.

Since
$$\frac{E}{B+R} = \frac{V}{R}$$

we have
$$\left.\begin{aligned}\frac{E}{B+1000} &= \frac{1\frac{3}{7}}{1000}\\ \frac{E}{B+100} &= \frac{1}{100}\end{aligned}\right\}$$

$$\therefore\ 1000E = \frac{10}{7}B + \frac{10{,}000}{7} \quad . \quad . \quad . \quad . \quad (1)$$

$$100E = B + 100 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

From (1)
$$100E = \frac{B}{7} + \frac{1000}{7}$$

Subtracting
$$\frac{6B}{7} = \frac{300}{7}$$

Substituting in (2)
$$\left.\begin{aligned}B &= 50 \text{ ohms.}\\ E &= 1{\cdot}5 \text{ volts.}\end{aligned}\right\}$$

CHAPTER IX

MEASUREMENTS INVOLVING STEADY CURRENTS

The Tangent Galvanometer—The Helmholtz Galvanometer—The Astatic Galvanometer—The Moving-Coil Galvanometer—The Moving-iron Galvanometer—Ammeters and Milliammeters—Voltmeters—The Resistance-box and the Rheostat—Reversing Switches—The Wheatstone Bridge—The Post Office Box—The Carey-Foster Bridge—The Potentiometer—Sensitivity and Resistance of Galvanometers—Other Experiments with Steady Currents.

The Tangent Galvanometer

THE Tangent Galvanometer (Fig. 53) consists simply of a number of accurately circular coils of wire arranged so that the plane of the coil is accurately in the magnetic meridian, and a magnetometer fixed horizontally, having a small magnetic needle with its centre at the centre of the coil. We may deduce the relation between the angular deflection θ of the magnetometer, and the current flowing through the coil, directly from the definition of unit current. For by definition (see p. 108):

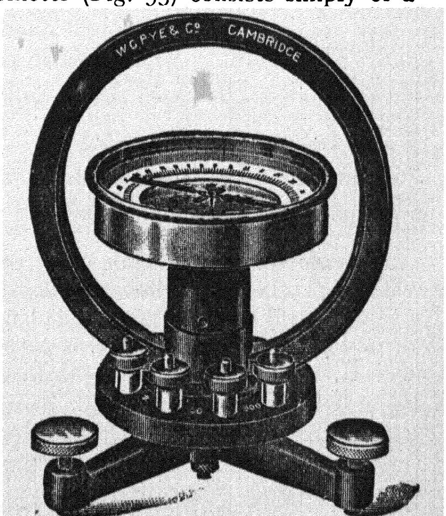


Fig. 53.

Fig. 53.

For current 1 abamp., length of wire 1 cm., radius of coil 1 cm., field at centre is 1 oersted.

For current 1 abamp., length of wire 1 turn, radius of coil 1 cm., field at centre is 2π oersted.

For current 1 abamp., length of wire n turns, radius of coil 1 cm., field at centre is $2\pi n$ oersted.

For current 1 abamp., length of wire n turns, radius of coil r cm., field at centre is $\frac{2\pi n}{r}$ oersted.

This change is difficult to follow, and not really sound, for it is the place where calculus is needed. For the correct proof see the chapters on Electromagnetism in Part II, pp. 335, 341. A rough outline of the argument which leads to it is that when we multiply the radius by r we multiply the length of wire producing the field by r; but we put every bit of it r times as far away, and by the inverse square law diminish the field in the ratio $\frac{1}{r^2}$. Thus at the same time we multiply the field by r and divide it by r^2. So we may suppose that the combined effect is to divide it by r, thus changing it from $2\pi n$ to $\frac{2\pi n}{r}$.

Continuing

Since 1 amp. = $\frac{1}{10}$ of an abamp.,

For current 1 amp., length of wire n turns, radius of coil r cm., field at centre is $\frac{2\pi n}{10r}$ oersted.

For current i amps., length of wire n turns, radius of coil r cm., field at centre is $\frac{2\pi ni}{10r}$ oersted.

The field due to the coil, being perpendicular to the coil which is parallel to H, the horizontal component of the earth's magnetic field, is itself perpendicular to the field. We have already seen that if a field H_1 is perpendicular to the earth's field H, the deflection θ of the magnetometer is given by $H_1 = H \tan \theta$.

Now $\qquad H_1$ in this case $= \frac{2\pi ni}{10r}$ oersted.

so $$\frac{2\pi ni}{10r} = H \tan \theta,$$

so $$i = \frac{10rH}{2\pi n} \tan \theta.$$

The quantity $\frac{10rH}{2\pi n}$, which depends on the galvanometer

and the earth's field, and is independent of i, is known as the Reduction Factor of the Galvanometer.

The Reduction Factor, R, may also be defined as the constant ratio $\frac{i}{\tan\theta}$, or the quantity by which $\tan\theta$ must be multiplied by, in order to find the current.

When the Tangent Galvanometer is used, it is most important to see that the magnetometer is accurately horizontal, and that the pointer by which readings are made on the scale is accurately perpendicular to the plane of the coil, and reading zero, when no current passes.

It is also important to see that no source of stray magnetic field, such as an ammeter containing permanent magnets, or a wire carrying a considerable current, or a piece of steel or iron, is near enough to affect the reading. It is possible to find out whether a particular object is causing interference by moving it, at the distance at which it lies, through at least a quarter of a circle, and seeing whether its motion produces any effect on the galvanometer deflection. Most Tangent Galvanometers give a choice of three numbers of coil-turns. The one in Fig. 53 has 2, 50, or 500 turns. The effective radius of each set of coils, and its resistance, should be marked on a label.

It is necessary always to use a Tangent Galvanometer with a reversing-switch, and the deflection should always be the mean of four observations; for each end of the needle should be observed for each direction of the current.

Values of θ less than 30° or greater than 60° should not be used, because a small error in the reading produces too large a percentage error in the result outside these values.

The Tangent Galvanometer is completely out of date for serious measurements. Its two main sources of error are the variations in H (which cannot be controlled or predicted) and the non-uniformity of the central field. This field was assumed to be uniform when the formula was calculated, but it is not. The Tangent Galvanometer is also relatively very insensitive. The advantages of the Tangent Galvanometer are that it measures current from first principles, using the definition quite directly, and this gives the learner a sound foundation to his ideas about currents; and that it is robust.

and cannot easily be damaged by rough use or moderately large currents. It is invaluable for the physicist or theoretical engineer to begin on, if he wants to be sound in the foundation of his subject, but is of little use for the practical engineer.

The Helmholtz Galvanometer

The Helmholtz Galvanometer (Fig. 54) is an improved Tangent Galvanometer, in which there are two coils, separated by a distance equal to the radius of each, with their planes parallel, and having the same number of turns.

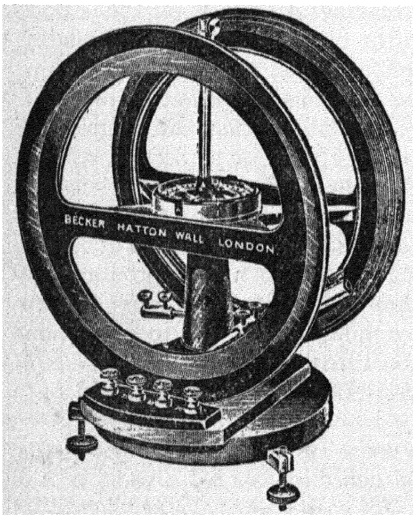


FIG 54.

The advantage of this arrangement is that the central field is much more nearly uniform. It can be shown that the intensity of the central field is $\frac{32\pi n i}{5\sqrt{5}r}$; and the proof of this relation appears as a question at the end of Part II.

The Astatic Galvanometer

The Astatic Galvanometer (Fig. 55) is another improved form of Tangent Galvanometer; but in this case the improvement is solely in sensitivity; and there is no exact mathematical relation between the deflection for a given current and the dimensions of the instrument.

The scheme for increasing the sensitivity is simple but effective. Where the Tangent Galvanometer has one freely suspended magnet inside the coil, the Astatic Galvanometer has two magnets one below the other, rigidly fixed and parallel to each other. They are magnetized in opposite directions, so that the controlling couple due to H is proportional to the

difference between their magnetic moments. By altering the magnetic moment of one of the magnets the controlling couple may thus be given any value, and it may be made very small by making the magnetic moments nearly equal.

Only one of the magnets, however, is inside the coil. The other is outside. The coil is made nearly flat instead of circular to allow this. Thus almost the whole effect of the couple due to the current acts on the inner magnet; and the couple on the outer magnet assists, instead of opposing, the couple on the inner one, since outside the direction of the field is reversed, as well as the direction in which the magnet is magnetized.

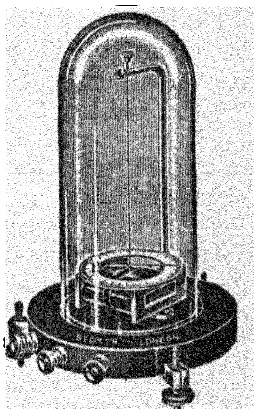


FIG. 55.

Hence it is roughly true to say that the couple due to the current depends on the sum of the magnetic moments of the suspended magnets, and that the controlling couple due to H depends on the difference between the magnetic moments.

It is thus approximately true to say that the sensitivity is proportional to $\frac{M_1 + M_2}{M_1 - M_2}$, and that it can thus be given any desired value.

This type of galvanometer is rather interesting theoretically, but it is troublesome to work, and not so good for any purpose as an ordinary moving-coil galvanometer.

The Moving-coil Galvanometer

It will be shown later, both in the chapter on Electromagnetism in Part I and in the chapters on Theoretical Electromagnetism in Part II, that, if a coil of wire carries a current in a magnetic field, a couple whose moment is

$$niHA \cos \theta \text{ units}$$

always acts on it, trying to turn it into a position in which it is perpendicular to the field, where n turns of wire, of area A

sq. cm., carry i absolute electromagnetic units of current, with the plane of the coil making an angle of θ with a field of H oersted.

Thus a coil of 1000 turns of area 2 sq. cm., in a field of 500 gauss, carrying a current of 1 milliamp., having its plane parallel to the field, would experience a couple of moment

$$1000 \times 10^{-4} \times 2 \times 500 \text{ units}$$

or

$$100 \text{ units.}$$

[Note that 1 milliamp., 10^{-3} amp., is put down as 10^{-4} unit of current because 1 amp. $= 10^{-1}$ absolute units; and that the unit of moment is the moment of a force of 1 dyne at a distance of 1 cm. from the fulcrum.]

The experimental work on rigidity also shows that the restoring couple on a twisted wire is proportional to the angle of twist.

Thus if we can succeed in suspending a coil in a uniform field radiating from its axis, from a wire which is twisted from its zero position when the coil experiences a couple due to a current, we have both

$$\text{Couple} \propto \theta, \text{ the angle of twist,}$$

and

$$\text{Couple} \propto niAH,$$

since if the field is radial the plane of the coil is always parallel to the field.

Thus, since n, A, and H are constant for the apparatus

$$i \propto \theta$$

or

$$i = k\theta,$$

where k is a constant of the particular galvanometer.

In the Moving-coil Galvanometer, therefore, the angle of deflection is directly proportional to the current causing it.

The radial field is arranged as shown in Fig. 56, in which N_1S_1 are the poles of a permanent magnet and N_2S_2 the induced poles of a short cylindrical piece of soft iron between N_1 and S_1. The field obviously is radial in the direction of the arrows.

The coil is arranged to surround the soft iron cylinder N_2S_2. It is suspended from a support vertically above the flat top of N_2S_2, and is kept in position by a very light spring fixed below it.

The arrangement of a typical Moving-coil Galvanometer is shown in Fig. 57, and an actual galvanometer is shown in Fig. 58. The suspension is usually of phosphor-bronze or quartz coated to make it conducting, since the current must reach the coil through the suspension.

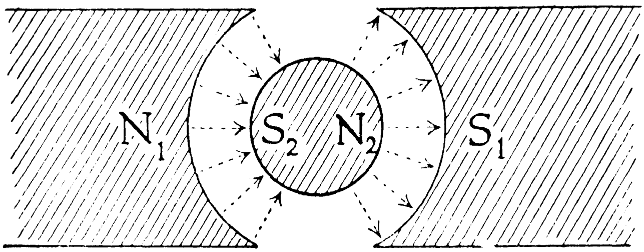


FIG. 56.

The Moving-coil Galvanometer, in one form or another, is more used than all other types of current-measuring instruments put together. It is the ordinary instrument in laboratories both for normal and high-sensitivity work, and is the moving part of all the best ammeters and voltmeters. A special form of it, the Ballistic Galvanometer, is used for the measurement of transient currents, and can be shown to give a deflection proportional to the total charge carried by such currents. It differs from the ordinary galvanometer only in having a very light moving system which responds very quickly to impulses.

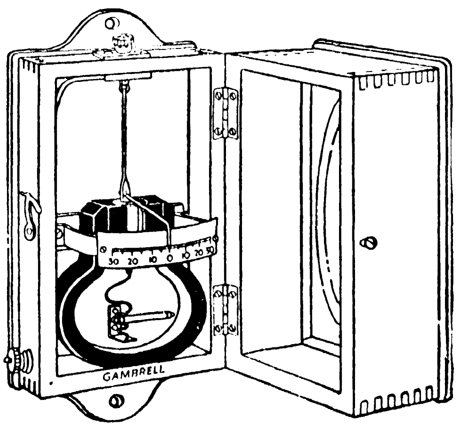


FIG. 57.

It cannot, however, be discussed without calculus. An

account of its use is given in Part II, and its detailed mathematical theory (which, though rather long, is quite easy if one does not allow oneself to be frightened by the look of it) is given in Ch. VI.

There are two principal ways of reading a Moving-coil Galvanometer.

It may be read by a pointer on a scale, as in Fig. 57. It may be read by observing a spot of light thrown on a scale from a mirror fixed to the suspension. Such a mirror is shown in Fig. 58, and the type of lamp and scale used is shown in Fig. 59. The scale is usually put at a standard distance of 1 metre,

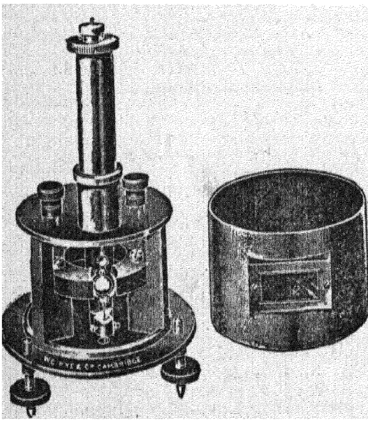


FIG. 58.

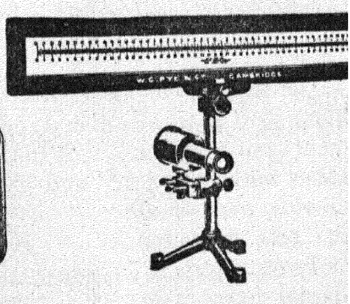


FIG. 59.

and the mirror is then a little concave mirror of radius of curvature 1 metre, so that a sharp image of the source is thrown on the scale.

It may be noticed that the deflection on the scale is proportional to $\tan\theta$, whereas the pointer-reading on a circular scale as in Fig. 57 is proportional to θ; so that it appears at first that the less sensitive pointer method of reading might give a more accurate result than the more sensitive lamp-and-scale method. This is theoretically correct, but in practice the error is small. The largest deflection used, at a distance of 100 cm., is 25 cm. from the zero position. Thus the maximum value of θ is $\tan^{-1}\frac{25}{100}$, which is $14^\circ\ 2'$. The value of

this angle in radians is 0·2450, so that the error introduced is just 0·5 in 25, or 2%. Thus when the deflection is 25 cm. on a straight scale it would have been 24·5 cm. on a circular scale. For a deflection of 10 cm. the corresponding error is 0·02 cm., which is negligible. For 15 cm. it is 0·11 cm., and for 20 cm. it is 0·26 cm. As the probable error of reading the scale is about 0·1 cm. the error is quite negligible for deflections up to about 12 cm. on either side (which gives a range of 24 cm really as one generally reverses the current), and for higher deflections one can allow for the error if need arises.

The Moving-iron Galvanometer

The Moving-iron Galvanometer serves the same purposes as the Moving-coil Galvanometer, but is in general inferior and cheaper. It forms the moving part of most cheap ammeters and voltmeters. There are, however, somewhat similar Moving-magnet Galvanometers which are more sensitive than Moving-coil Galvanometers. The general construction is on the same principle as the Tangent Galvanometer; but the earth's field is replaced by an approximately radial field set up by a permanent magnet. If the plane of a stationary coil coincides with the central plane of the field of the permanent magnet, and a small iron or magnet is freely suspended inside the coil, the field of the coil is approximately perpendicular to the field of the permanent magnet for small deflections.

This arrangement is in general cheaper and easier to make than that of the Moving Coil, because the difficult part of the arrangement of the latter is the fixing of a coil free to move and making good electrical connection. The suspension of a small magnet is obviously a much easier job.

The Moving-iron Galvanometer has, however, one very great advantage over the Moving Coil. Its sensitivity can be varied at will by having tappings on the coils at various points, and connecting these to external terminals as with the Tangent Galvanometer.

In the better instruments of this type, many small magnets are used instead of one, and the coils are of very special shape and position.

Moving-coil, Moving-iron, and Moving-magnet Galvanometers are discussed in greater detail in Part II.

Ammeters and Milliammeters

An ammeter or a milliammeter is simply a moving-coil or moving-iron galvanometer with a shunting resistance to give it the right sensitivity. A scale is attached to the instrument, and the shunt is adjusted to give the right readings on the scale. The difference between an ammeter and a milliammeter is only in the value of the shunt, the milliammeter shunt having, of course, a lower resistance relative to the galvanometer resistance.

The method by which the shunt-resistance is determined is best seen if examples are worked out.

There are two general methods of solving the problem. The common-sense method is given in Example 1, and the formula

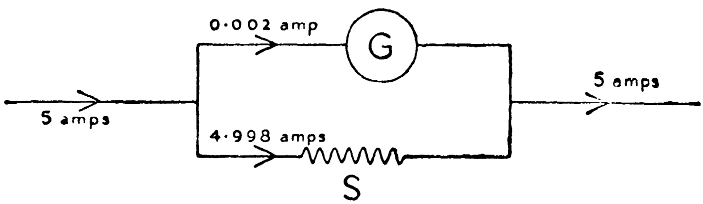


FIG. 60.

method in Example 2. A slightly harder problem is given in Example 3.

Example 1. Suppose you have a sensitive galvanometer that will only take 0·002 amp. as a maximum current, and you want to make it take just that amount out of a main current of 5 amps. what shunt must you put in parallel with the galvanometer, if its resistance is 40 ohms?

It is obvious that if 5 amps. are passing altogether, and 0·002 amp. is flowing through the galvanometer, then 4·998 amps. are flowing through the shunt S (Fig. 60).

It is also obvious, since G and S are in parallel, that the ratio of the currents through them is the ratio of their conductances.

$$\therefore \quad \frac{\frac{1}{G}}{\frac{1}{S}} = \frac{0{\cdot}002}{4{\cdot}998}$$

$$\therefore \quad \frac{S}{G} = \frac{1}{2499}$$

$$\therefore \quad S = \frac{40}{2499} \text{ ohm.}$$

Example 2 (Fig. 61). It is desired to send $\frac{1}{n}$ of the main current through a galvanometer of resistance G, shunted by a resistance S. Find the value of S, and the resistance of S and G in parallel.

(*a*) If i be the main current, then the galvanometer current is $\frac{i}{n}$, and consequently the shunt current is $\frac{n-1}{n}i$.

Since the currents are as the conductances,

$$\frac{\frac{1}{G}}{\frac{1}{S}} = \frac{\frac{i}{n}}{\frac{n-1}{n}i}$$

$$\therefore \frac{S}{G} = \frac{1}{n-1}$$

$$S = \frac{G}{(n-1)}.$$

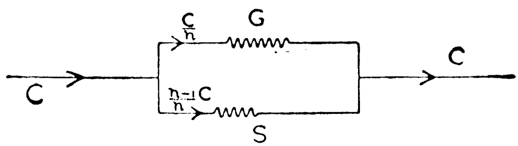


FIG. 61.

This gives a very simple rule. If $\frac{1}{n}$ of the main current is to be sent through the galvanometer, then the shunt resistance is $\frac{1}{n-1}$ of the galvanometer resistance.

(*b*) If R be the resistance of G and S in parallel, or, in other words, the total resistance of the ammeter formed by G and S in parallel, then

$$\frac{1}{R} = \frac{1}{G} + \frac{1}{S}$$

$$= \frac{1}{G} + \frac{n-1}{G} = \frac{n}{G}.$$

$$\therefore R = \frac{G}{n}.$$

Thus the resistance of an ammeter whose galvanometer carries $\frac{1}{n}$ of the main current is $\frac{1}{n}$ of the galvanometer's own resistance.

Example 3. The resistance of a galvanometer is 25 ohms, and its maximum current is 0·004 amp. Find what external resistance must be added to make it into a voltmeter reading to 2·5 volts.

If the same series resistance is used, but a shunt of 5 ohms is placed across the terminals of the galvanometer, what voltage will be needed to produce a full-scale deflection?

Total resistance of 2·5 volt-voltmeter $= \frac{2 \cdot 5}{0 \cdot 004} = 625$ ohms.

Hence the external series resistance is 600 ohms.

If a shunt of 5 ohms is in parallel with the galvanometer, then when the galvanometer carries 0·004 amp. the shunt will carry $\frac{\frac{1}{5}}{\frac{1}{25}} \times 0 \cdot 004$ amp., or 0·02 amp., and the total current will be 0·024 amp.

The total potential drop across series resistance and shunt will then be

$$\begin{aligned} &(600 \times 0 \cdot 024) + (5 \times 0 \cdot 02) \\ &= (14 \cdot 4 + 0 \cdot 1) \text{ volts} \\ &= 14 \cdot 5 \text{ volts.} \end{aligned}$$

Voltmeters

A Voltmeter is simply a galvanometer with a series resistance. The value of this resistance depends upon the current which gives full-scale deflection. Suppose for example 1 milliamp. gave full scale deflection. If the voltmeter were required to read 5 volts, it would require a total resistance of 5000 ohms, since 5 volts would send 1 milliamp. through this resistance. If the resistance of the galvanometer alone were 1000 ohms, an extra series resistance of 4000 ohms would be needed.

Example 4.—The maximum current taken by a sensitive galvanometer is 0·02 amp., and its resistance is 100 ohms. What series resistance must be added to make it into a voltmeter reading to 15 volts?

$$\text{Total resistance} = \frac{v}{i} = \frac{15}{0 \cdot 02} = 750 \text{ ohms.}$$

$$\begin{aligned} \therefore \text{ Series resistance needed} &= (750 - 100) \text{ ohms} \\ &= 650 \text{ ohms.} \end{aligned}$$

Example 5.—The galvanometer in Example 4 is shunted with a resistance of 10 ohms. What voltage is now required to cause full-scale deflection?

When the galvanometer carries 0·02 amp., the shunt carries $\frac{100}{10} \times 0{\cdot}02$, or 0·2 amp. So the galvanometer and shunt together carry 0·22 amp. The potential difference across them is $0{\cdot}2 \times 10$, or 2·0 volts.

The potential difference across the series resistance of 650 ohms is $650 \times 0{\cdot}22$, or 143 volts. So the total potential difference giving full-scale deflection is 145 volts.

Example 6.—In Example 5, what shunt across the galvanometer would have been needed to reduce the sensitivity to $\frac{1}{10}$ of its original value, or in other words to change the maximum reading from 15 volts to 150 volts?

Let this shunt be S.

Then the current through the shunt is $\frac{100}{S} \times 0{\cdot}02$, or $\frac{2}{S}$ amps.

The total current is $\left(\frac{2}{S} + 0{\cdot}02\right)$.

The potential difference across the galvanometer and shunt is $\left(\frac{2}{S} \times S\right)$, or 2 volts.

The potential difference across the 650 ohms outside is then

$$650\left(\frac{2}{S} + 0{\cdot}02\right).$$

So our equation is

$$650\left(\frac{2}{S} + 0{\cdot}02\right) + 2 = 150.$$

$$\frac{1300}{S} + 13 = 148$$

$$\frac{1300}{S} = 135$$

$$S = \frac{1300}{135} = 9\frac{17}{27} \text{ ohms.}$$

The Resistance Box and the Rheostat

The Resistance Box is a device for getting any one of a number of definite values of resistance at will. The commonest type (Fig. 62)[1] gives any whole number of ohms from 1 to 11,110. Inside the box are coils having resistances of 5000,

[1] The box shown gives 1,110 ohms only.

2000, 2000, 1000, 500, 200, 200, 100, 50, 20, 20, 10, 5, 2, 2, 1 ohms respectively. These are joined to large brass blocks

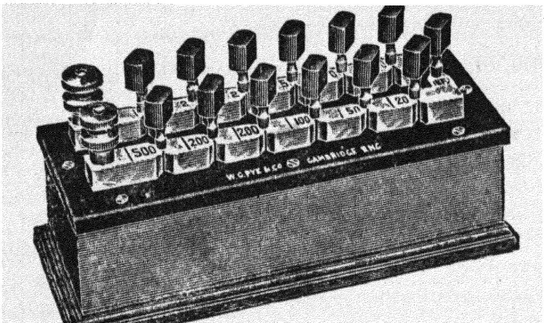


FIG. 62.

outside the box, and any one of these blocks may be directly connected to the next one (thus short-circuiting the resistance

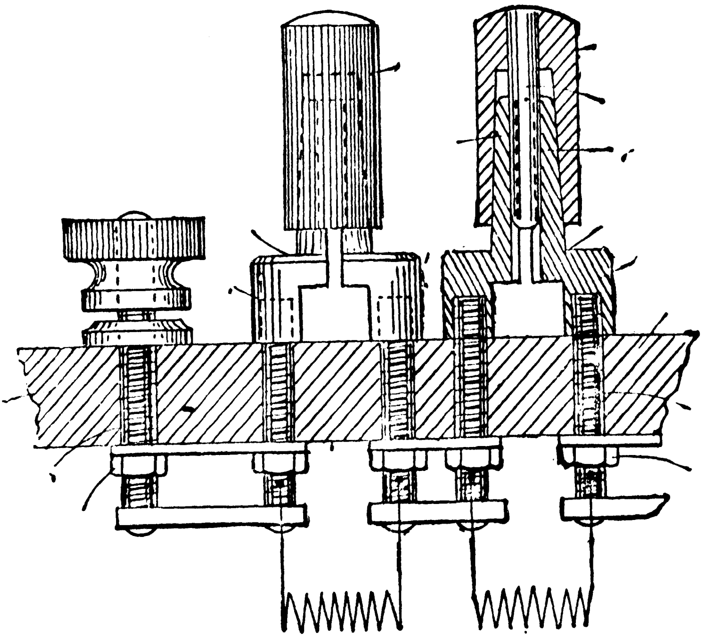


FIG. 63

joining them) with a plug which fits tightly between them (as in the box shown), or fits as a sleeve round two uprights (as in Fig. 63). The latter type is less likely to give a bad

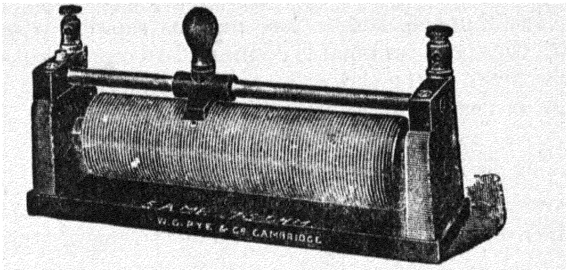


FIG. 64.

contact because of dirty surfaces, but the former is the type usually met. Either is perfectly satisfactory if it is kept clean. If all the plugs are in, the box should give a dead short-circuit from terminal to terminal, and any plug taken out introduces its associated resistance into the circuit. Any whole number between 1 and 11,110 may be obtained by taking some combination of the resistances given above. Only very small currents should be sent through resistance boxes.

The Rheostat is an arrangement giving a continuously variable resistance. A type carrying currents up to 5 amps. is shown in Fig. 64, and one carrying cur-

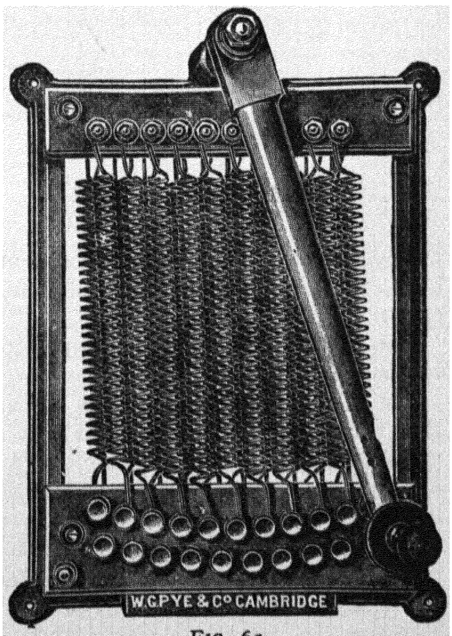


FIG. 65.

rents up to 50 amps. is shown in Fig. 65. Actually the resistance is not continuously variable in this case, but the instrument is called a rheostat nevertheless. The coils are in air in this type to allow them to get red hot without doing any damage, and to lose heat as rapidly as possible. In both types one terminal is connected to one end of the coil and the other to the slider; so that by varying the amount of wire in the circuit the resistance can be varied.

Reversing Switches

Reversing Switches, which allow the direction of flow of the current from a battery, or through a measuring instrument, to

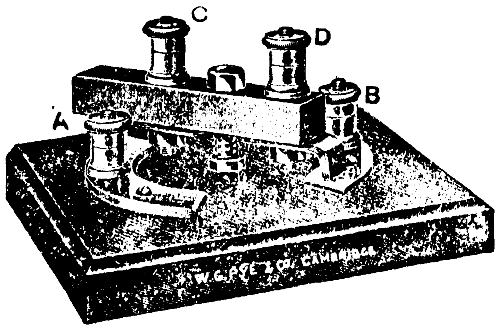


FIG. 66.

be reversed, are of the greatest value in cutting out zero errors in measuring instruments. They are of four main types, shown in Figs. 66 to 69.

The way of working a reversing switch is always best found by looking at it and thinking hard.

The method of working of the one in Fig. 66 is obvious. If the current flows in from the battery at A and out at B, the measuring instrument being across CD, then with the switch in one position the order of flow is ACDB, and in the other it is ADCB.

The rotor in Fig. 67 is made of ebonite, and holds two brass plates of peculiar shape, which connect B to C and A to D in one position, and B to D and A to C in the other. This switch

thus works like the one of Fig. 76, and the battery may be connected to A and B, or to C and D, but not on any account to A and C, or A and D, or B and D, or B and C, since any one of these connections would give a short-circuit in one position of the switch.

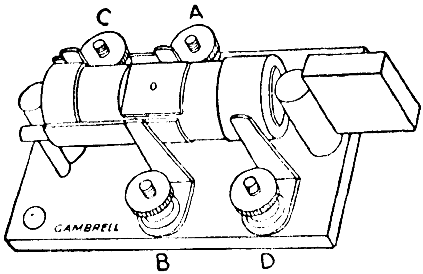


FIG. 67.

The switch in Fig. 68 is home-made. Holes ABCDEF are made in wood (or paraffin wax if a good insulator is needed), and they are filled with mercury. Connection is made by strips of copper from each hole to its own terminal, and

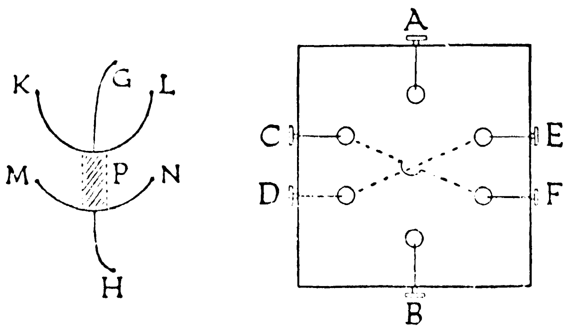


FIG. 68.

additional strips of copper connect CF and DE, but these two strips must cross each other in the middle without touching.

The rocker GHKLMNP completes the switch. G, L, K are bent pieces of thick copper or brass wire soldered together, and H, M, N are the same. Each set is fixed at its junction to P, a piece of ebonite.

G and H are placed permanently in the holes A and B, but they are longer than the other arms. Thus if the rocker lies over to the left, K is in C and M in D, and if it lies over to the right, L is in E and N in F. It is impossible for K, M, L, and N all to make contact together. Either K and M are in, or else L and N are in. The feature of this switch is that either terminals C and D are used, or else terminals E and F, but not on any account both. Let us suppose that the positive terminal of the battery goes to A, and its negative to B. Let the measuring instrument be across C and D. Then with the rocker to the left the current follows the course AC*D*B. With the rocker to the right it follows the course AE*DC*FB.

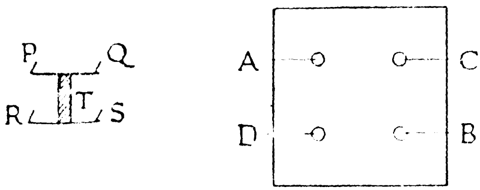


FIG. 69.

The fourth type (Fig. 69) is also home-made, and is the simplest, but has the disadvantage of having rather a slower change-over than the others.

Four holes ABCD, connected to terminals, form a square. They are again filled with mercury. The bridge-piece, PQRST, consists of two pieces of thick copper or brass wire PQ and RS, connected by an ebonite block T. PQRS form a square identical with the holes ACBD.

If the bridge-piece is fitted to the holes with P in A, Q in C, S in B, and R in D, then a current flowing in at A and out at B follows the path ACDB. If P is put in C, Q in B, S in D, R in A (the bridge-piece being thus turned through a right-angle) the current flows along the path ADCB.

The Wheatstone Bridge

The circuit-diagram (Fig. 70) shows the arrangement of a Wheatstone Bridge. P and Q are known fixed resistances. R is a variable resistance, and X the unknown resistance

which is being measured. A, B, C, D, are the points of junction of R and P, Q, and X, P and Q, and X and R respectively. V is a cell sending a steady current from A to B.

If the value of R is altered until no current nows through the galvanometer, then we know that the points C and D are at the same potential.

It therefore follows that the potential drop from A to C is the same as that from A to D.

If i_1 is the current flowing along ACB, and i_2 the current flowing along ADB, then

$$Pi_1 = Ri_2.$$

But in the same way the potential drop from C to B is equal to that from D to B, so

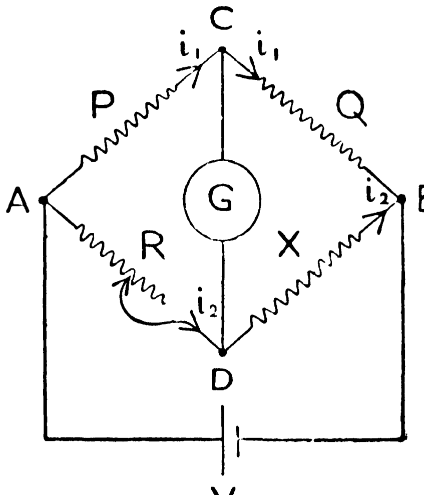


FIG. 70.

$$Qi_1 = Xi_2.$$

Dividing the first equation by the second we get

$$\frac{P}{Q} = \frac{R}{X};$$

and if P and X change places we get the equation in its usual form

$$\frac{X}{Q} = \frac{R}{P}.$$

The ordinary way of using the Wheatstone Bridge principle is in the metre-bridge (Figs. 71, 72). Here instead of having the resistances P and Q fixed and R variable, R is made fixed, and P and Q are the two parts of a metre-length of resistance-wire separated by the point of contact of T, the tapping-key which is connected to the galvanometer.

The ratio of the resistances P and Q is now, as for the potentiometer, simply the ratio of the lengths AT and TB. AC, EF, and DB are thick strips of brass or copper of negligibly small resistance.

V is an accumulator which gives a current through AB, and S is a series resistance to keep this current small.[1]

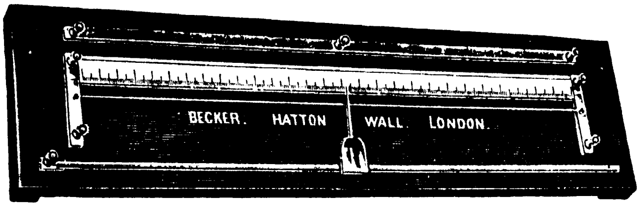


FIG. 71.

Two cross-checks of Wheatstone Bridge measurements are always necessary. The resistances R and X should be interchanged and fresh values of P and Q observed. If it is assumed that P always means the length opposite R, then the new values P′ and Q′ should be very near P and Q. If there

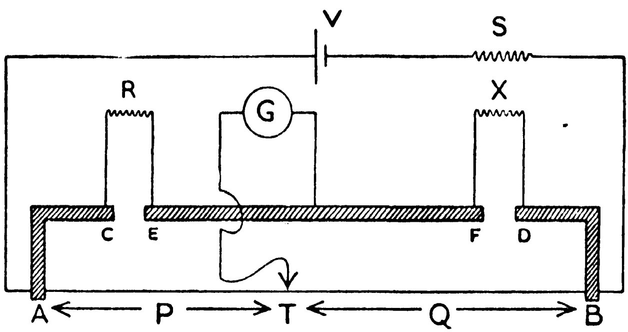


FIG. 72.

is more than about 1% difference between P and P′, there is a bad connection, or the leads to R or X have too much resistance.

Of course, the leads to R and X come into the equation; so when R and X are changed the leads should *not* be changed.

[1] It is the advantage of this method that variations in the E.M.F of V do not affect the result.

By making these leads as short and thick as possible, and by changing over R and X without changing the leads, we can reduce the error due to them until it is less than the error of reading.

The second cross-check consists in measuring X against a different (or, better, several different) known value of R. X and R should, of course, be exchanged for every value of R.

The values of R should be chosen so that contacts from which readings are taken always occur in the middle third of the wire, or at least between the 30 cm. and 70 cm. marks on a metre-scale. R should thus be between 2X and $\frac{X}{2}$. If it is outside this range the percentage error of the experiment is too large.

The management of the galvanometer presents a difficulty. Since it has to record only a zero deflection, a galvanometer of high sensitivity (about 1 scale division to a microamp. is a good value) may, and should, be used with due care. It is a good plan to shunt the galvanometer with a resistance-box, having *all* the plugs in the box for the preliminary search. This gives theoretically a dead short-circuit of the galvanometer; but actually, with a sensitive galvanometer, quite enough current gets through to show a deflection. An approximate point of contact can be found. Then the 1-ohm plug can be removed, and a closer position (to within about 2 mm. of the true one) can be found. Finally, the infinity plug is removed, and the final position is found with great accuracy. It is a good plan to use a Leclanché cell, rather than a 2-volt accumulator, to supply the current, because it cannot give too much. The resistance S need not be included if a Leclanché is used.

The making of the actual contact may also be rather a problem. If one uses a slider with a tapping-key it may be a nuisance to move about, and it may get pushed sideways so that its arrow gives a slightly wrong reading on the scale (but this error is more or less cut out when we change over R and X). If one uses a brass jockey, its contact edge must be fairly sharp to give a definite scale reading; and in the right position the uncertainty, whether there is no deflection because it *is* the right position or because the contact is bad, sometimes causes the enthusiast to press so hard upon the wire

that he makes a dent in it. Generations of enthusiasts produce a corrugated wire.

The ends of the wire, where it is fixed to the large brass contacts, are danger-points unless they are soldered, since in time partial oxidation (or general dirtiness) may produce bad contacts which are difficult to detect.

A typical calculation for a Wheatstone Bridge (in the Metre-Bridge form) follows :—

When $R = 2$ ohms, we have

$$P = 57{\cdot}65 \text{ cm. and } Q = 42{\cdot}35 \text{ cm.}$$
$$P' = 57{\cdot}31 \text{ cm. and } Q' = 42{\cdot}69 \text{ cm.}$$

Mean values $57{\cdot}48$ cm. and $42{\cdot}52$ cm.

$$\therefore \quad X = \frac{42{\cdot}52}{57{\cdot}48} \times 2 = 1{\cdot}479 \text{ ohms.}$$

When $R = 1$ ohm,

$$P = 40{\cdot}25 \text{ cm. and } Q = 59{\cdot}75 \text{ cm.}$$
$$P' = 39{\cdot}85 \text{ cm. and } Q' = 60{\cdot}15 \text{ cm.}$$

Mean values $40{\cdot}05$ cm. and $59{\cdot}95$ cm.

$$\therefore \quad X = \frac{59{\cdot}95}{40{\cdot}05} \times 1 = 1{\cdot}497 \text{ ohms.}$$

Mean value of $X = 1{\cdot}488$ ohms.

The Post-Office Box

The Post-Office Box was designed for the Post Office to find the position of a distant short-circuit in double telephone wires. Suppose the two legs of the wires are AB and CD. Let them touch at S. Then the resistance between A and C is (AS + SC), or 2AS. If the resistance of unit length of the wire is known, and the resistance of 2AS is observed, the exact position of S can be calculated.

The Post-Office box (Fig. 73) is the most accurate arrangement of the Wheatstone Bridge. It is simply a big resistance-box with special terminals arranged as shown in the diagrams (Figs. 74, 75). In this case the ratio arms, P and Q, can each have the values 1000, 100, or 10 ohms, but are not otherwise varied, and R may have any value from 1 ohm to 11,000 ohms, so that the value of the resistance to be measured can always be found to four significant figures.

Fig. 74 shows the way a Post-Office box is laid out, and Fig. 75 is a circuit-diagram with corresponding lettering.

AB and AD are the ratio-arms of resistance P and Q. BC is the arm of variable resistance R. The unknown resistance X is connected from the terminal C to the terminal D.

FIG. 73.

The switch S is connected to B inside the box, so the battery is connected between the terminal S and the terminal D

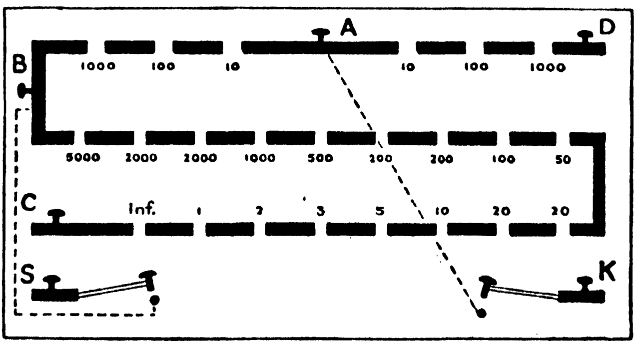


FIG. 74.

The switch K is connected to A inside the box, and the galvanometer is connected between the terminal K and the terminal C.

R is adjusted until there is no deflection of the galvanometer when both keys are pressed.

The key K should always be pressed last and let go first, for a reason that can be understood after the chapter on electromagnetic induction.

The galvanometer and its shunt should be treated as for the metre bridge. The reversal check cannot be applied at all and the check by the changing of R cannot be effectively applied, though it may be worth while to try a different ratio of P to Q. This only gives a check for large errors, and the value employing the largest value of R giving good sensitivity should be taken as final (except when we can change from $\frac{P}{Q} = \frac{100}{10}$ to $\frac{P}{Q} = \frac{1000}{100}$. Then we can average). There is, however, one useful check possible, if one possesses a reliable resistance-box, or one whose error at all values is known accurately enough. [A calibration slip is supplied by the makers with the best resistance boxes.]

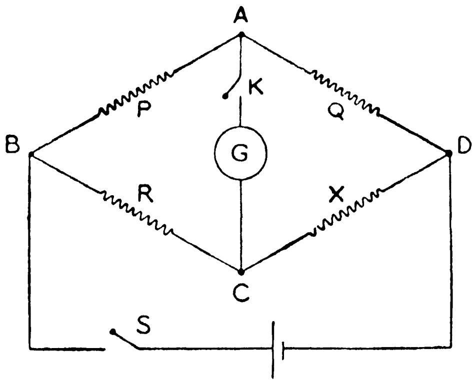


FIG. 75.

Suppose the galvanometer (unshunted) gives a deflection of 5 divisions left for R = 4253, and 15 divisions right for R = 4254, P and Q being 100 and 10 respectively, it is obvious that X is between 425·3 and 425·4. We may assume that X is (within the limits of error of experiment) on quarter of the way from 425·3 to 425·4. We may therefore take it as 425·325.

Let us now replace X by the resistance-box, giving it a value of 425; if we are using really thick short copper wire for

the leads to X, we can neglect its effect. If not, we must find its resistance from the book of constants and add this to 425. Let us then suppose that we believe the true value of the resistance of (box + leads) to be 425·05. If by measurement we get it to 425·10 by the method already described, it is obvious that in this neighbourhood the box is giving readings 0·05 ohm too high. The more correct value of X should then be taken as 425·275 ohms, since the box gave it as 425·325 ohms.

The Carey-Foster Bridge

This bridge, an application of the metre bridge, gives a good method of measuring small resistances quickly. It requires an accurate knowledge of the resistance per cm. length of the bridge-wire.[1]

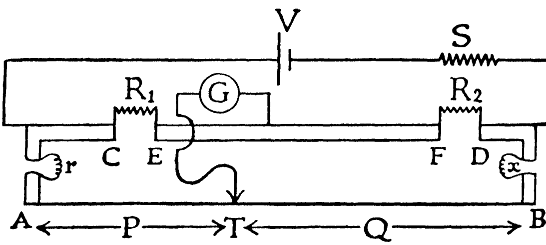


FIG. 76.

The circuit (Fig. 72) is the same as that of Fig. 76, except R and X are replaced by R_1 and R_2, two nearly equal resistances whose values need not be known; and that the brass connections are broken at A and B, and a small known resistance r, and a small unknown resistance x, are inserted at A and B respectively. [Most metre bridges are constructed with terminals and bridge-pieces allowing this to be done.]

Suppose a balance is obtained. Let P and Q be the resistances of the bridge-wire in ohms, instead of lengths of it in centimetres.

Then
$$\frac{R_1}{R_2} = \frac{P + r}{Q + x}$$
$$\therefore \frac{R_1}{R_1 + R_2} = \frac{P + r}{P + Q + r + x}$$

[1] This can be done by making r a small accurately known resistance (say $\frac{1}{2}$ ohm), and x a thick copper strip.

Now exchange r and x, and get a balance again at a point P′ from A and Q′ from B.

Similarly,
$$\frac{R_1}{R_1 + R_2} = \frac{P' + x}{P' + Q' + r + x}$$
$$= \frac{P' + x}{P + Q + r + x},$$

since $P' + Q'$ = total resistance of bridge-wire = $P + Q$.

It therefore follows that

$$P' + x = P + r$$
$$\therefore\ x - r = (P - P').$$

x thus differs from r by the resistance of the length of bridge-wire between the two points of contact.

By altering slightly the ratio of $\frac{R_1}{R_2}$, or by altering the value of r, any number of cross-checks can be obtained.

The Potentiometer

A Potentiometer (Fig. 77) is an instrument (or an arrangement of resistances) used for comparing potential differences. It consists of a long resistance wire through which a steady

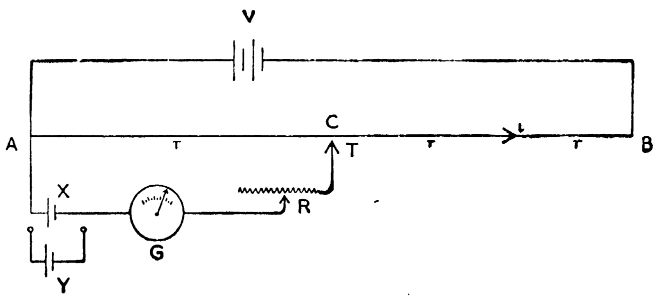


Fig. 77.

current is driven, and with which contact may be made at any point. (A common form of Potentiometer is shown in Fig. 78.)

An accumulator V drives a steady current from A to B along the uniform resistance wire AB. Contact may be made at any point C on the wire by means of a tapping-key T.

Suppose we want to compare the potential differences between the terminals of two cells X and Y.

Arrange X in series with a galvanometer G, and a large variable resistance R. If A is connected to the positive end of V, connect it also to the positive end of X. Connect the loose end of R to T.

Make the resistance of R as big as it can be (so as not to send too much current through G. R should be so big that the whole voltage available is not big enough to hurt the galvanometer) and move T about on the wire AB till a place C is found, such that no current flows through the galvanometer. Then gradually cut out the resistance R, and find the

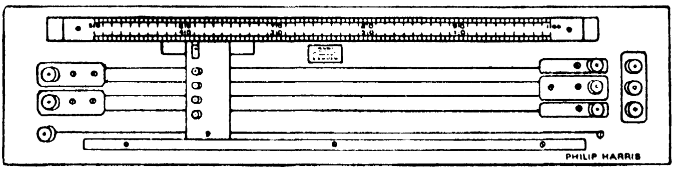


FIG. 78.

position of C as accurately as possible by using all the sensitiveness of the galvanometer.

If no current flows when contact is made, T must have been already at the same potential as C.

That is to say, the potential drop along AC is exactly equal to the potential difference across the cell X.

If a current i is flowing from A to B, and the resistance of AB is r ohms per cm., then the total potential drop from A to C is irAC.

So we have

$$X = ir\text{AC}.$$

If we now replace X by Y, and find contact is made at C_1, then

$$Y = ir\text{AC}_1.$$

$$\therefore \frac{Y}{X} = \frac{\text{AC}_1}{\text{AC}}.$$

So that the potential difference to be measured is proportional to the distance of the point of contact from A.

If accurate work with the potentiometer is required, a standard cell (described in the chapter on Cells) should be used. Special precautions should, however, be taken to see that no appreciable current is taken out of the cell by keeping R large until a good balance is obtained. R should, if possible, be variable between 100,000 ohms and zero.

As a cross-check on the ratio of the E.M.F.'s of two cells, the sum-and-difference method may be used.

The position of T giving no deflection is found first with the cells in series and their E.M.F.'s acting in the same direction; secondly, with the smaller of the two cells reversed. If x and y are the E.M.F.'s, then the ratio of the potentiometer readings is equal to $\frac{x+y}{x-y}$, and from this ratio $\frac{x}{y}$ can be calculated.

For example, in an experiment, contact for a Daniell cell was made at 306·0 cm., and for a Leclanché at 405·4 cm. For their sum it was at 710·0 cm., and for their difference at 100·0 cm.

Thus the direct value of

$$\frac{x}{y} = \frac{405 \cdot 4}{306 \cdot 0} = 1 \cdot 325.$$

By the sum-and-difference method

$$\begin{aligned} \frac{x+y}{x-y} &= \frac{710}{100} = 7 \cdot 1 \\ \therefore \quad x + y &= 7 \cdot 1x - 7 \cdot 1y \\ \therefore \quad 6 \cdot 1x &= 8 \cdot 1y \\ \frac{x}{y} &= \frac{81}{61} = 1 \cdot 328. \end{aligned}$$

The mean of these values should be taken.

For accurate work, a high-resistance Potentiometer, whose resistance can be very accurately known, should be used. Such a Potentiometer is shown in Fig. 79.

Sensitivity and Resistance of Galvanometers

The sensitivity of a galvanometer has several possible meanings. Moreover, any galvanometer may be regarded as having a voltage-sensitivity and a current-sensitivity.

The current-sensitivity may mean :—

(*a*) The scale deflection produced by unit current.

(*b*) The reciprocal of the current required to produce 1 mm. deflection on a scale 1 metre away.

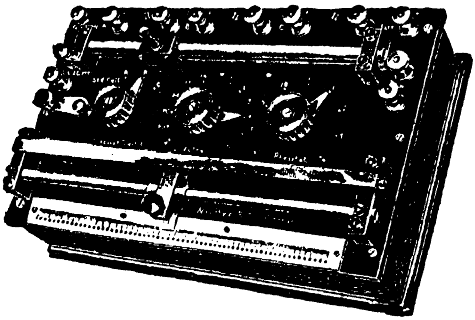


FIG. 79.

(*c*) The current required to produce this deflection. [This is a bad definition, since the numerical value of it decreases as the galvanometer becomes more sensitive.]

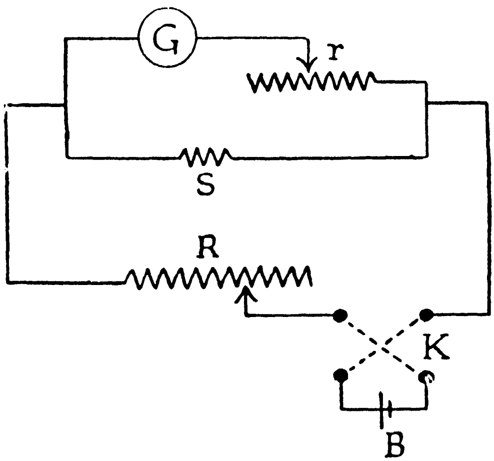


FIG. 80.

(*d*) (and the most modern and useful). The deflection in millimetres, on a scale 1 metre away, produced by a current of 1 microampere (10^{-6} ampere).

We shall use the latter definition.

The voltage-sensitivity is then the deflection per microvolt. It is obviously equal to the current-sensitivity divided by the resistance.

The resistance of the galvanometer must, of course, be found before its sensitivity. It may be roughly determined by sending enough current through the galvanometer, when shunted by a small resistance, to produce full-scale deflection, and then finding what resistance in series with the galvanometer will halve this deflection. Fig. 80 shows a circuit for this method.

B is a cell or battery, K a reversing-key, R and *r* resistance boxes, S a shunting resistance of about 0·1 ohm (its exact value need not be known), and G is the galvanometer.

All the resistance is cut out of *r*. R is then gradually reduced from its maximum value (which should be not less than 10,000 ohms) until G shows full deflection; *r* is then increased, leaving R untouched, until G shows exactly half deflection. As the current in the main circuit is effectively independent of the value of *r*, we have effectively doubled the resistance of the galvanometer when we halved its deflection. Thus G may be taken as equal to the final value of *r*.

This method, though useful as a beginning or a cross-check, is only approximate. When G is known roughly, the galvanometer should be clamped, and G should be measured, like an ordinary resistance, by using a Post-Office box. Great care must be taken not to send enough current through the galvanometer to heat up its coils appreciably.

When the resistance is accurately known, the measurement of sensitivity is perfectly straightforward. The galvanometer is connected in series with a resistance box, and an accurately-known potential difference is applied across the combination.

The practical difficulty arises because a good galvanometer may have a sensitivity of the order of 10,000 mm. per microamp. Full-scale deflection of 250 mm. is thus produced by a current of $\frac{1}{40}$ microamps., or $2{\cdot}5 \times 10^{-8}$ amps.

The ordinary accurate resistance box has a resistance of

about 10,000 ohms. So the required potential difference should not exceed $2 \cdot 5 \times 10^{-8} \times 10^{4}$, or $2 \cdot 5 \times 10^{-4}$ volts.

It would be possible to get round this, theoretically, by using a shunt of known value, with the circuit of Fig. 80, and adjusting R and r as required. But it would be difficult for this shunt to be known accurately, because leads and bad contacts might introduce an enormous proportionate error.

One can, however, get the required potential difference by

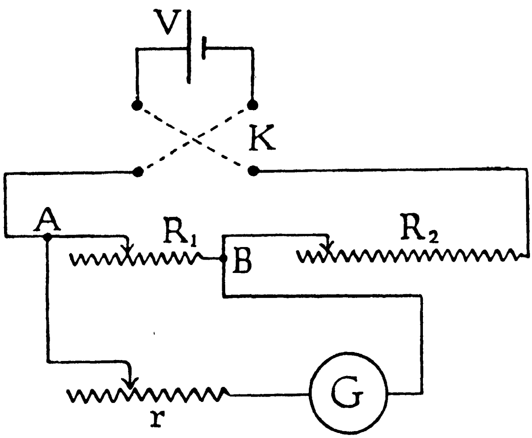


FIG. 81.

a potentiometer arrangement, as shown in Fig. 81, which requires three resistance boxes R_1, R_2, and r.

If V is a cell of known E.M.F. and negligible internal resistance, the potential difference between the points A and B is clearly $\frac{R_1 V}{R_1 + R_2}$, and this can be given any required value.

Let us consider an actual experiment.

The E.M.F. of the accumulator V, which had been recently charged, was found by potentiometer measurement, from comparison with a standard Weston cell of E.M.F. 1·0183 volts, to be 2·080 volts.

The resistance of the galvanometer was 824 ohms, and a maximum current of about 2×10^{-8} amps. was expected. If a resistance of 9176 ohms from an 11,110 ohm resistance

box were put in series, the total resistance of $(G + r)$ would be 10,000 ohms, which would make the arithmetic easy. A potential difference of $2 \times 10^{-8} \times 10{,}000$, or 2×10^{-4} volts would be required to send this current.

This quantity is $\frac{1}{10000}$ (approximately) of the E.M.F. of the cell. So R_1 should be about $\frac{1}{10000}$ of the value of $(R_1 + R_2.)$

R_1 and R_2 being both similar to r, having a maximum of 11,110 volts each (though R_1 would actually have been equally useful with a small maximum value, such as 20, 50, or 100 ohms, provided it remained accurate), R_1 was given the value 1 ohm, and R_2 the value 10,399 ohms. Thus $(R_1 + R_2)$ was 10,400 ohms, and the P.D. across R_1 was $\frac{1}{10{,}400} \times 2{\cdot}08$, or 2×10^{-4} volts, as was required.

The reduction of potential across R_1, because a 10,000 ohm resistance $(G + r)$ was in parallel with it, was quite negligible.

This gave, in fact, a mean deflection (both left and right deflections were taken, by using K_1 for every reading) of 235 mm.

R_1 was then replaced by a stretched resistance-wire of exactly 1 ohm total resistance, and the value of the deflection was observed for various positions on this wire, the potentials of which with respect to the zero end could be easily calculated from the length of wire to the point of contact.

The current through G was then $\frac{1}{10{,}000}$ of this potential, since the resistance of the G circuit was 10,000 ohms. A graph of deflection against current was thus plotted, and the best straight line drawn through it. The first point found came accurately on this line, so the sensitivity was 235 mm. for $2 \times 10^{-4} \times 10^{-4}$ amps., or 11,750 mm. per microamp.

Other Experiments with Steady Currents

The experiments described—the Metre Bridge, Post-Office Box, Carey-Foster Bridge, Potentiometer, and Sensitivity of Galvanometer—are the most important to understand as a beginning, but there is a large choice of experiments to be found in intermediate or advanced practical text-books. Everyone should look up and perform three or four at least. Such experiments are :—

The Ayrton and Mather Universal Shunt for Galvanometers.
Kelvin's method of determining Galvanometer resistance.
Mance's method of determining Battery resistance.
The Kelvin Bridge for small resistances.
Measurement of Current by Potentiometer.
Internal Resistance of Cell by Potentiometer.

Experiments on the variation of Resistance with Temperature, and on the Thermo-electric Effect, will be described in the next chapter.

CHAPTER X

HEATING EFFECTS

The Volt—Power and the Watt—Simple Problems on Heating Effects and Power Supply—Callendar and Barnes' Experiment—Variation of Resistance with Temperature—Measurement of Temperature Coefficient of Resistance—The Thermo-electric Effect—Measurement of E.M.F.'s of Thermo-couples—Thermopiles and Bolometers—Pyro-electricity and Piezo-electricity.

The Volt

THE Potential Difference in Volts between two points A and B in a circuit may be regarded as the rate of production of energy, in the form of heat, in joules per coulomb, or joules per ampere per second.

If we say, for example, that the voltage of direct-current mains is 110 volts, we really mean that each ampere of current flowing produces 110 joules of energy per second. Thus the rate of production of heat by a current is clearly given by

$$H = ViT \text{ joules,}$$

where H is measured in joules, V in volts, i in amps., and T in seconds.

Since by Joule's experiment J joules of energy are equal to 1 calorie of heat, we may write

$$H \text{ calories} = \frac{ViT}{J} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The mean value of J is usually taken as 4·182. For practical purposes we may take it as 4·2.

If the current supplying the heat is flowing through a resistance of R ohms, then we have $V = iR$, and $i = \frac{V}{R}$.

By making these substitutions we get

$$H = \frac{i^2RT}{J} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$$H = \frac{V^2T}{RJ} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

One is sometimes asked for the Laws of Production of Heat in an Electric Circuit. There is none.

The First Law of Thermodynamics, that heat and energy are equivalent, holds here as elsewhere. It is the foundation of Joule's determination that 1 calorie of heat equals J joules of energy. If the Laws of Production of Heat in an Electric Circuit are asked for, the equations (1), (2), and (3) above probably form the answer required, though they may perhaps be better stated in words. An alternative way of stating the " Laws " is :—

For a given current, the rate of production of heat is directly proportional to V and R. [Equations (1) and (2).]

For a given potential difference, the rate of production of heat is directly proportional to i and inversely proportional to R.

For a given resistance, the rate of production of heat is directly proportional to V^2, and directly proportional to i^2.

These statements are not laws, because the meaning of "volt" makes it inevitable that they should be true if Joule's relation holds.

Power and the Watt

Power is defined as the " Rate of Working," and the watt, the practical unit of power, is a rate of working of 1 joule per second. Since the number of joules done per second by a current i, flowing across a potential difference V, is Vi, it follows that the rate of working, W, in watts, of a current i flowing across a P.D. of V, through a resistance R, is given by

$$W = Vi$$

or

$$W = \frac{V^2}{R}$$

or

$$W = i^2R.$$

Similarly, from our equations for the rate of heat production in the last section, we see that

$$H = \frac{WT}{J},$$

where H is in calories and T in seconds.

A subsidiary unit, generally used commercially, is the kilowatt, a rate of working of 1000 watts.

The Board of Trade Unit of Energy (sometimes sold as the Unit of Electricity) is the kilowatt-hour, the work done in an hour at the rate of 1 kilowatt. This 1 kilowatt-hour = 3,600,000 joules, or about $3{\cdot}6 \times 10^{13}$ ergs.

As a unit costs about 3*d*. (though its price varies with the district), an erg is probably about the cheapest thing one could buy, for a single one (though it is difficult to get them singly) would cost about $2 \times 10^{-14}d$. Other useful relations are—

$$\begin{aligned} 1 \text{ joule} &= 0{\cdot}737 \text{ ft.-lb.} \\ 1 \text{ H.P.} &= 746 \text{ watts.} \\ 1 \text{ kw.} &= 1{\cdot}34 \text{ H.P.} \end{aligned}$$

Simple Problems on Heating Effects and Power Supply

Many simple problems merely require the relations stated in the last section between V, i, R, and W. Electric light bulbs are generally listed by their voltage and wattage, the first telling what mains they may be used on, and the second their rate of consumption of power or production of heat.

To find the current through, and the resistance of, a 200-volt 50-watt lamp we have

$$i = \frac{W}{V} = \frac{50}{200} = \tfrac{1}{4} \text{ amp.}$$

$$R = \frac{V^2}{W} = \frac{200 \times 200}{50} = 800 \text{ ohms.}$$

Such a lamp is working at a rate of $\frac{50}{1000}$ kw., and would thus use $\frac{1}{20}$ of a unit in an hour.

Other simple problems follow :—

Example 1.—An electric kettle, made of copper of specific heat 0·1, weighs 4000 gm. and holds 2 litres of water. It takes 4 amperes at 200 volts. How long would it take to boil, starting at 20° C.? [J = 4·2.]

Thermal capacity of kettle = 2000 + (4000 × 0·1) calories = 2400 calories.

Total number of joules required = 2400 × (100 − 20) × 4·2.

Number of joules supplied per second = 4 × 200.

$$\text{Time taken in minutes} = \frac{2400 \times 80 \times 4{\cdot}2}{800 \times 60}$$
$$= 16{\cdot}8 \text{ minutes.}$$

Example 2.—A motor, working a crane, runs at 60% efficiency for 10 hours, taking 500 amps. at 200 volts. It is lifting cargo through an average height of 40 ft. from the hold of a ship on to the quay. How many tons of cargo does it lift?

We should get the same result if the motor worked for 6 hours at 100% efficiency. Thus the useful work done is $\frac{500 \times 200}{1000} \times 6$, or 600 kilowatt-hours. This is $1 \cdot 34 \times 600$, or 804 horse-power-hours.

Since 1 H.P. = 33,000 ft.-lb. per minute, number of tons of

$$\text{cargo lifted through 40 ft.} = \frac{804 \times 33{,}000 \times 60}{2240 \times 40}$$
$$= 17{,}759 \text{ tons.}$$

At 3*d*. a unit, the energy required for this work would have cost 3×600*d*., or £7 10*s*.

Since 1800*d*. are required to raise about 18,000 tons through 40 ft., a rough value for energy is about 400 foot-tons a penny.

An ordinary large electric stove, taking 20 amps. at 200 volts, consumes 4 units an hour, and therefore consumes one pennyworth in five minutes. Thus the heat produced by an electric stove of this size in five minutes requires about 400 foot-tons of work to produce it.

Example 3.—A wire of circular section is attached to an electric supply of constant voltage. The wire is then removed, and redrawn so that its total volume remains the same but its length is doubled. Compare the rates of production of heat in the wire before and after these operations, assuming that its specific resistance remains unaltered.

Since $R = \frac{SL}{A}$, and L is doubled while A is halved (because the volume LA remains the same), R is multiplied by 4. But the voltage remains the same. Thus the rate of production of heat is divided by 4.

Example 4.—A wire, carrying a certain current, settles at a steady temperature of 200° C., its surrounding temperature being 20° C. The rate of loss of heat from its surface obeys Newton's law of cooling (is proportional to the difference of temperature between itself and its surroundings), and its temperature coefficient of resistance is 0·004 at 200° C.

The current is increased by 1%.

Find the new steady temperature of the wire.

When the temperature is steady, the rate of loss of heat is equal to the rate of supply.

Thus we have—

$$\frac{\text{New rate of supply of energy}}{\text{Old rate of supply of energy}} = \frac{\text{new rate of loss of heat}}{\text{old rate of loss of heat}}.$$

Let the temperature rise from 200° C. to $(200 + t)$° C., while the current increases from i to $1{\cdot}01\,i$.

Then, assuming Newton's Law of cooling, we have

$$\frac{(1{\cdot}01)^2 i^2 R(1 + {\cdot}004t)}{i^2 R} = \frac{(200 + t) - 20}{200 - 20}.$$

Taking $(1{\cdot}01)^2 = 1{\cdot}02$

we have $1{\cdot}02(1 + {\cdot}004t) = \dfrac{180 + t}{180}$,

from which we get $t = 14$° C. rise.

Final temperature = 214° C.

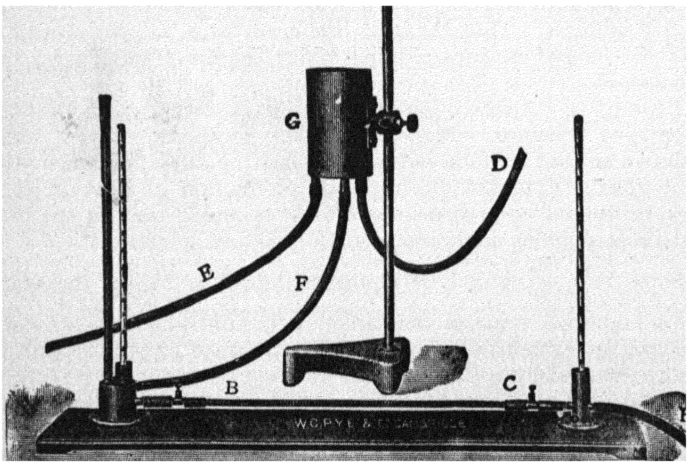


FIG. 82.

Callendar and Barnes' Experiment [1]

This experiment was devised to determine J electrically.

A continuous stream of water is made to flow through a tube containing the spiral wire BC (Figs. 82 and 83).

[1] This is a very good method. There is no need to find a water equivalent and no lag in thermometer readings.

The rate of flow is controlled in an ingenious way by the cistern G. Water flows into the bottom by the tube D, out of the bottom past the spiral of wire by the tube F, and out to the sink by the overflow tube E, whose opening inside the cistern is near the top.

The overflow tube E keeps the top level of the water in the cistern constant, and the difference between this level and that of the tube containing the spiral may be adjusted by raising or lowering the cistern on the stand. The rate of flow of the water depends on this difference of level, and it may therefore be kept constant at any desired value within a fairly wide range.

The rate of flow is easily measured by catching the water

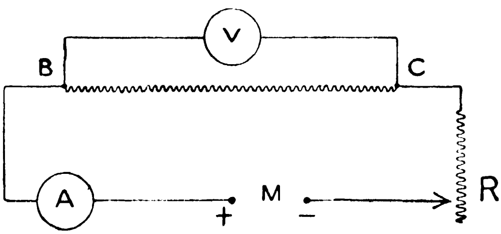


FIG. 83

from the outflow tube K in a measuring glass, and finding how much flows in a given time.

The apparatus is connected up as shown in the circuit-diagram (Fig. 83), A being an ammeter in series with BC, and V a voltmeter in parallel with BC[1]. M are the mains, and R is a series resistance (usually consisting of the carbon lamps of a charging-board for charging accumulators), by which the P.D. across the spiral, or the current through it, may be controlled.

There should also be several spirals of different resistances for comparison.

The difference of temperature ($T_2 - T_1$) on the thermometers, multiplied by the number of grammes of water flowing per second, gives the rate of supply of heat, by the spiral carrying the current, in calories per second.

Since the current, P.D., and resistance can all be varied and measured, the laws of production of heat can all be easily and

[1] The original experimenters used specially accurate methods, nearer to first principles, for measuring voltage and current.

fairly accurately illustrated. This apparatus cannot, of course, test the production of heat by the flow of currents through electrolytes or gases.

Variation of Resistance with Temperature

The resistance of metals increases with rise of temperature. The increase of resistance is approximately proportional to the rise of temperature, provided the range is not too large. Thus for ranges near 0° C. we may write

$$R_t = R_0(1 + \alpha t)$$

where R_t is the resistance at $t°$ C., R_0 at 0° C., t the temperature centigrade, and α a constant depending on the material of the wire.

The constant α is called the temperature coefficient of resistance. A few interesting values of α, with the temperature-range, are given.

Substance.	Temperature Range.	α.
Copper	Near 18° C.	0·0043
Iron	Near 18° C.	0·0062
Steel	Near 18° C.	varies from 0·0016 to 0·0042
Platinum	− 100° C. to + 100°C.	0·00365
Mercury	0– 24° C.	0·0009
Silver	0–100° C.	0·0040
Tungsten	0–100° C.	0·0051
Nichrome	Near 20° C.	0·00017
Brass	Near 18° C.	0·0010
German-silver	Near 18° C.	from 0·00023 to 0·0006
Eureka Alloy (Constantan)	Near 18° C.	from − 0·00004 to + 0·00001
Manganin	Near 20° C.	from 0·000002 to 0·00005

This table shows why manganin or eureka wire is generally used for resistance boxes. Manganin, an alloy of copper and manganese, with in some cases a little iron, nickel and silicon, is good. This alloy has a maximum specific resistance in the neighbourhood of 25° C. for an alloy with 83 parts copper, 14 parts manganese, 2·5 parts nickel, and 1·5 parts iron (not percentages; 101 parts in all). The temperature of maximum resistance varies with the composition. The best alloy for resistance boxes appears to be 87·02% copper, 9·13% manganese, 3·56% nickel, 0·19% iron, 0·10% silicon.

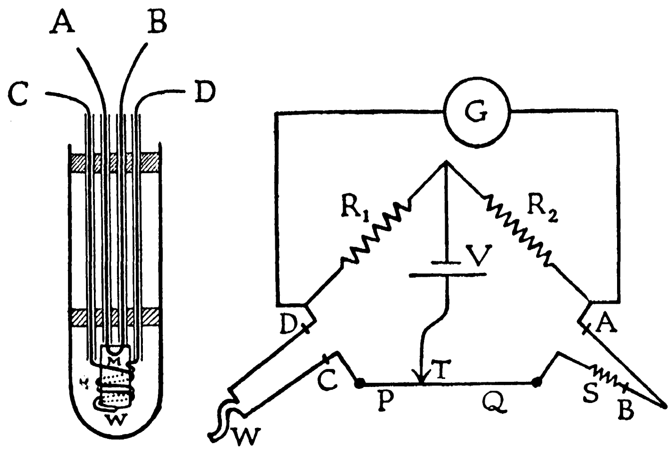


FIG. 84.

Measurement of Temperature Coefficient of Resistance

Fig. 84 shows the arrangement.

A, B, C, D are copper leads of the same cross-section, held in position as shown by glass tubes. The glass tubes are held in the large boiling tube shown by two pierced corks which fit the tube tightly. A thermometer (omitted to keep the diagram simple) is also held in holes in the corks. It is fixed so that its bulb is as near as possible to W, the wire to be tested. W has its ends soldered to the ends of the wires C and D, and it is wound non-inductively on M, a piece of mica. To wind a wire non-inductively, make it return on itself, so

that the magnetic field due to one half of it cancels out that due to the other. The wire AB is continuous through the two central glass tubes.

The rest of the circuit is connected as in the right-hand part of the figure, which gives the theoretical circuit. V is a cell, R_1 and R_2 nearly equal resistances (or quite equal resistances) of about the same value as W (and 1 ohm is a convenient value for R_1, R_2, and W; the wire must be thin if it is to be reasonably short. S is a resistance box with values ranging from 0·1 ohm to 10 ohms. P and Q are the two parts of a metre or half-metre bridge wire on which a jockey T makes contact. The copper wires A, B, C, D, coming from the glass tubes, are connected at the points A, B, C, D in the circuit. W is the wire under test.

It is most important in this apparatus that all connections should be sound, and that the resistance of all leads not included in the calculation should be negligible. Wide strip copper should be used for the leads if possible.

First the tube containing the wire under test is surrounded by melting ice. After about 20 minutes it should have reached a temperature of 0° C. T is then put in the middle of the bridge-wire, and S is adjusted till it is as near as possible to the value giving an exact balance. The final adjustment for exact balance is made with T.

It is important to realize that zero deflection for balance is only obtained *immediately* after contact. If contact is made for an appreciable time, the heating effect in the wire W will send up the temperature and cause a deflection. It is for this reason vital to have no inductive resistances whatever in the circuit; for an inductive resistance would produce a kick at the moment of contact—the only moment that matters.

Let us suppose that R_1 and R_2 are identically equal. Let r_0 be the resistance of W at 0° C., r_t at t° C. Then when the balance is obtained with W at 0° C.,

$$r_0 + P + \text{resistance of leads CD} = S + Q + \text{resistance of leads AB}.$$

By the construction of the apparatus, the resistances of AB and CD are equal; and, moreover, since they lie side by side, and are similar in length, cross-section, and material, their resistances remain equal whatever the temperature-

effects in the apparatus, since such temperature-effects must be the same for both.

Thus $$r_0 = (S_0 + Q_0 - P_0).$$

The bridge wire is calibrated, so P and Q are known. If the temperature of W rises to t, a fresh adjustment of S and T is made, and again

$$r_t = (S_t + Q_t - P_t).$$

Thus r_t can be observed for any value of t.

The Thermo-electric Effect

When two dissimilar metals are in contact, it is (rather surprisingly) found that they are at different potentials. The potential difference between the pairs of metals varies with temperature. It is about 0·75 volt for copper and zinc at ordinary temperatures. If a complete circuit is made with two metals, and the two junctions between the metals are maintained at different temperatures, an E.M.F., called the Thermo-electric E.M.F., acts round the circuit, and causes a current to flow.

The contact potential difference at a single junction is discussed in more detail in the chapter on Cells.

Such an arrangement as this is called a Thermo-couple, and the whole effect is the Thermo-electric effect, or the Seebeck effect. The energy of the current comes from the excess of the heat supplied at the hot junction over that removed at the cold junction.

As might be expected, the inverse effect is observed. If a current is made to traverse a circuit of two different metals, the temperature rises at one junction and falls at the other. This is nothing to do with the ordinary heating effect of a current, because it depends on the direction of the current. Thus, for temperatures round about 50° C., for example, antimony is positive to bismuth, so that in the thermo-electric circuit the current flows from bismuth to antimony at the hot junction.[1] An artificial current sent in this direction causes cooling at the junction; and if it flows in the opposite direction it causes heating.

[1] The following is a useful mnemonic. "Current flows from *A*ntimony to *B*ismuth across the *C*old junction."

This effect—the heating or cooling of the junction by the passage of a current—is called the Peltier effect.

The Peltier effect may be observed with simple apparatus. The two junctions are put in the two bulbs of a Differential Air Thermometer. A current is driven through the circuit. The mercury-thread moves toward the junction which is being cooled—that is, the junction across which the current flows from Bismuth to Antimony. This method neatly avoids the masking of the Peltier effect by the Joule effect.

It can now be seen that the effect of a thermo-electric current itself is to cool its hot junction and to heat its cold junction, thus trying to restore equilibrium.

A third related effect is the Kelvin, or Thomson, effect, which Lord Kelvin discovered in 1851. Heat is absorbed or liberated when a current flows in an unequally heated conductor. In iron heat is liberated when the current flows from cold to hot, and absorbed when it flows from hot to cold. In copper the effect is the reverse of this.

Though over small ranges the Seebeck E.M.F. is proportional to the difference of temperature between the junctions, it is found that the direction of the E.M.F. can be changed if the hot junction is raised to a high enough temperature. This "thermo-electric inversion" occurs at about 550° C. for a copper–iron junction. Iron is positive to copper below this temperature, and copper is positive to iron above it.

A few values of the Seebeck E.M.F., for circuits in which one junction is at a temperature of 0° C. and the other at 100° C., are given below.

The direction of flow of the current is from negative to positive across the hot junction. The pairs of metals are given in the order negative–positive; the order showing the direction of flow of the current across the hot junction.

Bismuth–Constantan	0·00306 volt.
Constantan–Nickel	0·00180 ,,
Nickel–German silver	0·00064 ,,
German silver–Palladium	0·00044 ,,
Palladium–Lead	0·00097 ,,
Lead–Manganin	0·00016 ,,
Manganin–Copper	0·00017 ,,
Copper–Iron	0·00086 ,,
Iron-Antimony	0·00310 ,,

These values are cumulative. Thus the E.M.F. for a Bismuth–Nickel circuit is 0·00306 + 0·00180, or 0·00486. Similarly, the Bismuth–Antimony circuit gives a value which is the sum of all those shown. It will be found by addition to be 0·01120 volt.

At temperatures between 0° and 100° we may thus make a list of metals in the order of either vertical column in our list, in which any metal higher in the list is negative to one lower.

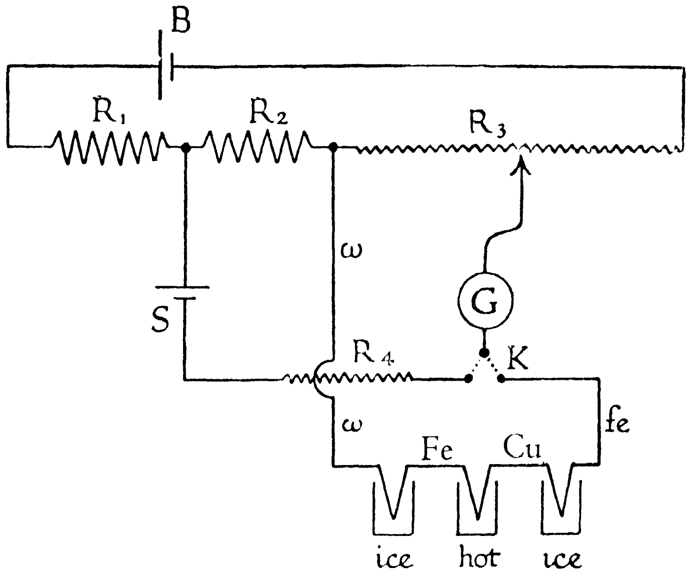


FIG. 85.

Gold and zinc almost coincide with copper; and magnesium, brass and tin nearly coincide with lead.

Measurement of E.M.F.'s of Thermo-couples

The circuit is shown in Fig. 85.

B is an accumulator whose E.M.F. has settled down to a steady value. A steady value is best obtained by discharging a fully charged accumulator at the required rate for at least half an hour. R_1 and R_2 are 11,110 ohm resistance

boxes, R_3 is the wire of a metre (or better 2-metre) bridge. Its resistance has been accurately determined in a separate experiment. G is a galvanometer, S a standard cell such as a Weston cell. R_4 is a large resistance, in series with the standard cell, to protect it from having current taken out of it. About 10,000 ohms is a good value for R_4.

Three thermo-couples are used, two ice-cold and one hot, as shown. In the figure a Cu-Fe couple is considered. The junctions should be made by welding the two wires, and one wire should be enclosed in a glass tube, right up to the junction, to prevent contact between them anywhere else. [If welding is too difficult, and great accuracy is not needed, fairly satisfactory results can be obtained by cleaning both wires thoroughly with emery-paper and twisting them together.] The couple should be inside a small test-tube with a thermometer. The test-tube can stand in a beaker containing melting ice for the cold junctions and hot water for the hot junctions. K is a two-way key.

It might be supposed that junctions between dissimilar metals in other parts of the circuit would affect the situation. They may, if the temperature is uneven over the apparatus, or if they are not adequately shielded from the ice and hot water of the junctions under observation. But in general stray effects cancel out, because mercifully every finite piece of wire has exactly two ends, and the algebraic sum of the E.M.F.'s of any number of junctions at the same temperature is automatically zero.

First the temperature difference is adjusted to have its maximum value, and contact is made on R_3 at a point near R_2. The thermo-couples are put in the circuit and S is cut out. A value of $(R_1 + R_2)$ is found which gives an approximate balance. Let us suppose that the value of R_3 is 2·5 ohms, and that contact is made at a point $\frac{1}{2}$ an ohm from R_2. If a Cu-Fe couple is used with a temperature-difference of 0° C. to 100° C., the E.M.F. is of the order of 0·001 volt. Thus if the E.M.F. of the accumulator B is about 2 volts (as it should be), $(R_1 + R_2)$ should be $\left(\frac{2}{0{\cdot}001} \times 2\right)$ ohms, or 4000 ohms.

This gives us a convenient value to work with, and merely serves the purpose of allowing us to use the whole, or

nearly the whole, of R_3, and thus get as much accuracy as possible.

We now begin the exact measurement. R_1 and R_2 are each made 2000 ohms.

The switch K is put to the left, and a careful adjustment is made to get a balance with the standard cell S. This is best done by making the connection to R_3 in the centre of the range, and making small adjustments to R_1 and R_2 so that their sum remains constant. The final adjustment is made on R_3.

Thus the value in ohms (equal to $R_2 +$ the part of R_3 between R_3 and the contact) of the resistance, across which the fall of potential is 1·0183 volts, is found. 1·0183 volts is, of course, the E.M.F. of the Weston cell. Corrections for room temperature (according to the book of constants) are made if necessary.

K is now put over to the right, and R_3 is adjusted to find the balance. Since the number of volts per ohm is known from the first adjustment, the E.M.F. of the Thermo-couple can be easily calculated.

As many different temperatures as time allows should be obtained, and a graph should be plotted showing the variation of E.M.F. with temperature. Over such a small range as 100° C. this graph is in general nearly a straight line.

Thermopiles and Bolometers

A Thermopile is an arrangement for measuring the rate at which energy is received by a number of thermo-couples in series. A relation between the rate at which radiant energy is being received, and the thermo-electric current produced, is obtained by calibration.

Much can be learnt from this; particularly the temperature of a source of known area and distance from which only feeble radiation can be received, such as a star.

The total energy radiated per square cm. per sec. from a source at temperature $T_1°$ Absolute, to a receiver at $T_2°$ Absolute, has been found to be proportional to $(T_1^4 - T_2^4)$. This is known as Stefan's Law.

Thus we may say

$$W = \sigma(T_1^4 - T_2^4)$$

where W is the rate of emission of energy in ergs per square centimetre per second, and σ, known as Stefan's Constant, has a value whose determinations range from $5{\cdot}45 \times 10^{-5}$ to $5{\cdot}89 \times 10^{-5}$.

From the rate of reception of energy the rate of emission of energy can be calculated, if the distance and area of (or solid angle subtended by) the source are known.

The principle of the Bolometer is similar, but it consists of a piece of resistance wire forming one arm of a bridge. When it receives energy its temperature rises. Consequently its resistance rises, and the balance of the bridge is destroyed. A current therefore flows through the galvanometer, and a relation between this current and the rate of reception of energy can be found by calibration.

One of the methods of calibrating a Bolometer is as follows. The Bolometer wire is made one arm of a Wheatstone bridge in which the temperature-coefficient of resistance is negligible in the other three arms. The bridge is balanced for a particular value of the main current. The main current is then altered, and the resistance of the Bolometer wire rises with rise of temperature, while that of the other arms remains almost unchanged. The additional energy absorbed by the Bolometer wire can be calculated, and the change of resistance observed. Thus a relation between absorption of energy and change of resistance of the Bolometer wire can be obtained.

Pyro-electricity and Piezo-electricity

Unequal-axis crystals and hemi-morphic equal-axis crystals are electrically polarized by heating or cooling (pyro-electricity). The direction of polarization due to cooling is opposite to that due to heating. Similar effects result from compression (piezo-electricity). The pyro-electric effect is due to distortion caused by change of temperature, and so is really piezo-electric. The total polarization varies as the total force applied. This effect is reversible. An applied voltage between opposite outer faces of the crystal, coated with conducting material, causes mechanical stress in the crystal, and hence slight deformation. As a crystal slab has a natural period of vibration, resonance, with a large peak current for a fixed applied alternating voltage, can be obtained at the crystal's natural frequency. This fact is useful if one wants oscillations of reliable frequency. Various crystals, such as quartz or Rochelle salt, are used for this purpose.

CHAPTER XI

ELECTROLYSIS

Faraday's Laws—The Ionic Theory—Evidence for the Ionic Theory—The Copper Voltameter—The Silver Voltameter—The Water Voltameter—Back E.M.F. in Voltameters—Secondary Actions—Practical Applications.

Faraday's Laws

CURRENTS through conducting liquids (other than molten metals) differ from other currents in causing chemical decomposition. Such a process of chemical decomposition is called *electrolysis*, and the substance decomposed is called the *electrolyte*. The metallic plates dipping into the liquid, by which the current enters and leaves, are called *electrodes*. The electrode at higher (or more positive) potential is the *anode*, and that at lower (or more negative) potential is the *cathode*.

Typical electrolytes are acids (such as H_2SO_4), bases (such as NaOH), and salts (such as $CuSO_4$). For electrolysis they are usually dissolved in water, or, less frequently, they may be electrolysed in the fused state.

In general constituents of the acid or salt appear at the electrodes when the current passes, but in some cases secondary actions occur, depending on the nature of the electrode, and on the primary product of electrolysis. The electrolysis of hydrochloric acid gives an example of the theoretically simple case, in which no secondary actions occur. If a current is passed between carbon electrodes immersed in concentrated HCl, bubbles of hydrogen appear at the cathode, and bubbles of chlorine at the anode. If these gases are collected separately with precautions as to solubility, it is found that the volumes (at the same pressure) are equal. By Avogadro's hypothesis it follows that the numbers of molecules of each gas liberated are equal. The liberation of the constituents of the acid is the only chemical action connected with the process.

The electrolysis of copper sulphate[1] gives the primary action and one secondary action. If a current is passed between copper electrodes in copper sulphate solution, copper atoms both leave the anode and reach the cathode. The primary action is the deposition of copper from the solution on the cathode. The secondary action is the recombination of SO_4 from the solution with the copper of the anode. These processes are explained in detail in the next section.

Faraday's Laws of Electrolysis, enunciated in 1833, state:

1. The mass of any substance liberated at an electrode is proportional to the total quantity of charge passing.

2. The mass of any substance liberated by a given quantity of charge is proportional to the chemical equivalent of the substance.

The mass of a substance liberated by the passage of 1 coulomb is called the *Electrochemical Equivalent* of the substance. The *Chemical Equivalent* of an element is the ratio of the mass of the element to the mass of the hydrogen it replaces, or combines with, in chemical combination. It is thus numerically equal to the mass in grammes of the element which replaces, or combines with, one gramme of hydrogen. This definition is the original one. When we want a numerical definition, we may take the Chemical Equivalent as being the Atomic Weight of the metal divided by its Valency. This result follows directly from the chemical definition of valency.

The Ionic Theory

Faraday's laws were based on experiment; their theoretical interpretation, which developed after they were formulated, is known as the Ionic Theory.

In order to explain the Ionic Theory, we must give a brief account of the theory of Electrolytic Dissociation.

Solutions are homogeneous mixtures of two or more substances. The constituents of solutions are not in chemical

[1] This example gives an ideally simple case of primary and secondary actions, and represents the view ordinarily advanced in elementary text-books. It is useful in giving a clear idea of the meaning of terms, but may not be true, for the SO_4 ions may never react at all. In this case there would merely be two primary actions, deposition of Cu ions at the cathode and emission at the anode, and equilibrium would be maintained as explained on p. 211.

combination (though there may be partial chemical interaction in certain circumstances). Many types of solution exist. There are solutions of gases in gases; of gases, liquids, or solids, in liquids; and of solids in solids. Three of these classes are of interest in Electrolysis: those of liquids in liquids, and of solids in liquids, and those of gases in liquids.

In solutions containing two substances, there is usually much more, both by volume and by weight, of one substance present than of the other.

The component present in greater proportion is called the *solvent*; the other, the *solute*. The solute is said to be *dissolved* in the solvent. The solute is evenly distributed throughout the solvent. Some properties of solutions depend upon the ratio of the number of molecules of the solute to the number of molecules of the solvent.

For example, if any substance is dissolved in water, the freezing-point of the solution is invariably lower than the freezing-point of the pure water. The depression of the freezing-point is found to be directly proportional to the number of molecules of the solute present in a given mass of the water.

Now liquid solutions are of two classes: those which can carry an electric current (called "electrolytes"), and those which cannot carry a current.

In an electrolytic solution the depression of the freezing-point is always greater than the theoretical depression calculated from the number of molecules of the solute present.

In order to account for this fact, it is assumed that when an acid, a base, or a salt is in solution in water its molecules are broken up into parts so that the number of separate particles of the solute is greater than the number of molecules of the solute.

For example,

$$\begin{array}{lcl} H_2SO_4 & \rightarrow & 2H^{\cdot} + SO_4'' \\ NaCl & \rightarrow & Na^{\cdot} + Cl' \\ NaOH & \rightarrow & Na^{\cdot} + OH' \\ CuSO_4 & \rightarrow & Cu^{\cdot\cdot} + SO_4'' \end{array}$$

In this notation dots represent positive charges, and dashes represent negative charges. The sum of the charges on the ions resulting from the break-up of one molecule must, of course, be zero, since the molecule as a whole was uncharged.

Any charged particle is called an *ion*.

A positively-charged particle is called a *cation* because it goes to the cathode.

A negatively-charged ion is called an *anion* because it goes to the anode.

In order to explain the nature and size of ionic charges, we must consider experiments in an entirely different range of Physics.

In 1895, Sir J. J. Thomson, then Cavendish Professor of Experimental Physics in the University of Cambridge, and afterwards Master of Trinity, was investigating the properties of the rays which left the cathode in a discharge-tube.

The rays were allowed to pass through an electric field of intensity E, and a magnetic field of intensity H at right angles to E. Since the rays were deflected by both these fields they were evidently charged particles. If the particles had a charge e, the force on each due to E was Ee; and if they were moving with velocity v, the force due to the magnetic field was Hev. Since the magnetic deflection was perpendicular to H, the two deflections could be made to oppose one another in the same plane. When they were so adjusted that the particles were undeflected, the two forces were equal, so that

$$\mathrm{E}e = \mathrm{H}ev,$$

and thus

$$v = \frac{\mathrm{E}}{\mathrm{H}}.$$

If now the magnetic field alone was allowed to act, the particles described a circle of radius r which could be shown to be given by

$$r = \frac{mv}{\mathrm{H}e}$$

where m was the mass of a particle.

In this equation H, E, and r could be measured, and thus from the equation

$$\frac{e}{m} = \frac{\mathrm{E}}{\mathrm{H}^2 r}$$

e/m could be determined.

The value of $\frac{e}{m}$ for all the particles without exception was found to be

$$1{\cdot}77 \times 10^7 \text{ e.m.u. per gm.}$$

By an independent series of experiments the charge on the charged particles was measured.

If a saturated vapour is suddenly subjected to an expansion, its temperature falls and it condenses into small drops. Such drops condense more readily on charged than on uncharged particles.

If, therefore, ions are present when the condensation takes place, the drops condense on the ions.

If no electric field is present, the cloud settles down under the action of gravity.

Sir G. G. Stokes showed that if a small drop of density ρ and radius r is falling through a gas of viscosity η under the action of gravity g, then

$$v = \frac{2}{9}\frac{g\rho r^2}{\eta}$$

where v is the terminal velocity of the drop (or the rate at which the cloud could be observed to settle down).

Everything in this equation could be observed except r. From the value of r so obtained, the weight of the drop could be calculated.

An electric field was now applied acting vertically upwards. When this field E was enough to hold the cloud, of drops of weight w, stationary, then

$$Ee = w.$$

This gave a direct value for e, the charge on the drop.

This method was repeated even more effectively by R. A. Millikan, of Chicago University, who used an oil spray. He observed individual oil drops in the field of a microscope.

It was found that in all cases the charge on each drop had the value

$$e = 1{\cdot}591 \times 10^{-20} \text{ e.m.u.}$$

or some small exact multiple of it.

A mass of evidence, which cannot be given here, pointed to the conclusion that this charge was the charge on the cathode particles, for which

$$\frac{e}{m} = 1{\cdot}77 \times 10^{7} \text{ e.m.u.}$$

These particles were thus seen to be uniform particles of mass

$\frac{1{\cdot}591 \times 10^{-20}}{1{\cdot}77 \times 10^{7}}$, or 9×10^{-28} gm. The cathode particles thus appeared to be the ultimate atoms of negative electricity, and they were called *Electrons.*

A great number of quite independent methods gave (within the limits of experimental error) the same values for e and $\frac{e}{m}$.

A very concise and clear account of the work on the atomic nature of electricity is given in J. A. Crowther's "Ions, Electrons, and Ionizing Radiations."

Now the general evidence from electrolysis indicates that all ions of the same kind carry the same charge; and the measured value of this charge is invariably found to be equal, within the limits of experimental error, to the electronic charge or some multiple of it.

Let us now return to the behaviour of those solutions which are electrolytes.

Where well-understood disturbing effects, such as polarization, can be eliminated, electrolytes obey Ohm's Law, and any applied E.M.F., no matter how small, suffices to maintain a current through them.

For this reason, and from the evidence of the depression of the freezing-point of electrolytes, and from other related phenomena, it is supposed that the solution is ionized *before the potential difference is applied.* The solution is not ionized by the current, since there is no sign that the current has to do the work which would be required for this purpose. The ions are already in existence, ready to carry the current. As to the percentage of molecules of the solute which are dissociated, there is considerable discussion.

From two points of view it would appear that the ionization of the solute is never complete, but approaches completeness as infinite dilution is approached.

In order to consider these points of view we must consider a new term—*Equivalent Conductivity.* It is the conductivity, measured in reciprocal ohms, of a solution containing 1 gram-molecule (58·5 gm. for NaCl) of the solute when placed between electrodes 1 cm. apart. One would naturally expect that the molecular conductivity would increase in proportion to the number of ions present to carry the current.

* 1949 values are $\frac{1{\cdot}602 \times 10^{-20}}{1{\cdot}759 \times 10^{7}} = 9{\cdot}1 \times 10^{-28}$ gm.

In the following table column 1 gives the concentration of the solution of NaCl in water in gram-molecules of NaCl per litre of water.

Column 2 gives the percentage additional depression of the freezing-point above that due to the number of molecules of NaCl present.

Column 3 gives the molecular conductivity in reciprocal ohms.

1.	2.	3.
1·0	67%	74
0·1	80%	92
0·01	93%	102
0·001	98%	106·5
0·0001	99%	108

It is obvious that the depression of the freezing-point is approaching 100% of that due to the number of NaCl molecules present, and that the molecular conductivity is approaching a maximum value as the dilution approaches infinity.

It would therefore appear that the dissociation of the NaCl into $Na^{\cdot}$ ions and Cl' ions only approaches completeness as the dilution approaches infinity. Sutherland, however, suggested in 1907 that dissociation is complete at all concentrations, and that the variation in columns 2 and 3 is due to change in the (certainly very large) forces between the ions. These forces would naturally vary with the average distance apart of the ions, and would decrease as concentration decreases. As a result of further work by Debye and Hückel (1923) and Onsager (1926), Sutherland's view is now generally accepted.

The lowering of conductivities and freezing-points below the expected values is explained by assuming that pairs of ions, though dissociated, still maintain a loose relationship which prevents them from being effective current-carriers.

The Ionic Theory is also confronted with another difficulty — the source of the energy required to produce dissociation.

Work must be done to separate the two oppositely-charged parts of a molecule. What is the source of the energy which supplies this work?

The fact of the separation is easily explained in a general way. The dielectric constant of water is very high—about 81—and it must thus reduce 81-fold the electrical forces binding the parts of the molecule together, since E, the electric intensity between two charges Q, distant r from each other, is given by

$$E = \frac{Q^2}{kr^2}$$

in a medium of dielectric constant k.

It does not, however, appear to be invariably true that of two solvents that with the largest dielectric constant necessarily gives the greatest molecular conductivity for a given concentration.

In order to account for the energy of dissociation, we must presumably suppose that the solvent is polarized, and stores energy as energy is stored in the dielectric between the plates of a charged condenser.

It is difficult to get independent evidence for this assumption. The absence of a satisfactory solution for the energy problem is the chief difficulty in the Ionic Theory.

There is evidence about those solutes, such as sodium chloride, which form crystals.

W. H. and W. L. Bragg found a method of investigating the structure of crystals with X-rays. Crystals may be regarded as forming solid diffraction-gratings with the spacing required to form diffraction-patterns from X-rays. From these diffraction-patterns the arrangement of the particles in the crystal may be deduced.

It is found in sodium chloride crystals that the molecules are already dissociated in the crystal, although no solvent is present. It is built of charged atoms of sodium and chlorine. These observations evidently have considerable bearing on the problem of dissociation in electrolytes. This method,

however, can tell us nothing about substances which do not form crystals.

There is one more aspect of the problem of electrolytic dissociation to be considered.

However far apart the electrodes may be placed, the products of electrolysis appear simultaneously and immediately at both electrodes as soon as the potential difference is applied. These products cannot have come from the same molecules, since the rate of migration of the ions, found by the work of Hittorf and Kohlrausch, is much too small. The fastest ion, the hydrogen ion, is found to travel at about 10 cm. per hour when the electric field is 1 volt per cm.

This fact gives independent evidence that dissociation exists all through the electrolyte before the current begins to flow.

The essential features of the Ionic Theory may now be rapidly summarized.

Ions are always present in an electrolyte.

They carry one or more electronic charges, positive or negative.

The positive, or electro-positive, ion, also called the cation, migrates toward the cathode when a potential difference exists between the electrodes; and the negative, or electro-negative, ion, also called the anion, migrates toward the anode.

Typical cations are the hydrogen ion, $H^{\cdot}$, in acids; and the metallic radicle in salts, such as $Cu^{\cdot\cdot}$ in $CuSO_4$.

Typical anions are the acid radicle, such as SO_4'' from H_2SO_4; the OH' ion from bases such as NaOH; and the acid radicle, SO_4'', from salts such as $CuSO_4$.

We will now consider in detail how the Ionic Theory accounts for Faraday's Laws.

Suppose now that an atom of mass m units of atomic weight has a valency n, and the electronic charge is e coulombs. The chemical equivalent is, by definition, $\frac{m}{n}$.

The number of atoms reaching the cathode (taking the case of an atom like copper which is positively charged in solution) while 1 coulomb passes is $\frac{1}{ne}$, since 1 coulomb is unity in our measure of quantity, and each atom carries a charge of ne coulombs.

Thus the electrochemical equivalent is

$$\frac{m}{ne}$$

since $\frac{1}{ne}$ atoms arrive, each of mass m.

Since this is the mass per coulomb, the theory fits the first law automatically; and since $\frac{m}{ne}$ is proportional to $\frac{m}{n}$ (e being the electronic charge, a universal constant), the second law is a consequence of the theory.

Let us, from the accepted values of the fundamental constants, calculate the values of the electrochemical equivalents of hydrogen, copper, and silver.

The charge on an electron (according to the revised value, 1948, due to spectroscopic and other evidence) is $1 \cdot 602 \times 10^{-20}$ e.m.units. or $1 \cdot 602 \times 10^{-19}$ coulombs.

The atomic weight of hydrogen may be taken as 1·00814 (proton 1·0076, electron 0·000054) on the atomic weight scale on which oxygen is 16·00. The atomic weight of copper is 63·57, of silver 107·88.

Some ions, such as those of hydrogen and silver, carry 1 electronic charge, and others, such as those of copper, carry 2 electronic charges.

For sulphuric acid, we may write

$$H_2SO_4 \longrightarrow 2H^{\cdot} + SO_4''.$$

For HCl,

$$HCl \longrightarrow H^{\cdot} + Cl'.$$

But for copper

$$CuSO_4 \longrightarrow Cu^{\cdot\cdot} + SO_4''.$$

This breaking-up is considered to be due to the high dielectric constant (about 80 for water) of the solvent, which enormously reduces the strength of the electrostatic bonds holding the component parts of the molecule together.

The valency of an atom or group is the number of hydrogen atoms with which it can combine, or which it can replace. Since hydrogen has one electron per atom, a substance like chlorine, which takes up 1 electron extra, has

valency 1; and the SO_4 radicle, which can take up 2 electrons extra, has a valency 2.

Also copper can replace two hydrogen atoms (compare $CuSO_4$ and H_2SO_4), and its valency is also 2; and hydrogen has unit valency by definition. In considering electrolysis, it does not matter whether the valency is positive or negative.

It is thus clear that an ion in electrolysis carries a number of electronic charges (positive or negative) equal to its valency. SO_4 carries 2 electrons extra, chlorine 1 electron extra, hydrogen 1 electron short, and copper 2 electrons short.

In order to work out the electrochemical equivalent, we have to know the true mass in grammes of one unit on the atomic weight scale on which oxygen is 16·00.

This has been determined through Dr. F. W. Aston's Mass Spectrograph method, and the value at present (1948) accepted is

$$1{\cdot}660 \times 10^{-24} \text{ gm.}$$

The mass m of an individual atom is then $1{\cdot}660 \times 10^{-24} \times$ its atomic weight, and if n is the valency (1 for hydrogen or silver, 2 for copper) and e the electronic charge, the formula $\frac{m}{ne}$ gives us for hydrogen

$$\frac{1{\cdot}660 \times 10^{-24} \times 1{\cdot}00814}{1 \times 1{\cdot}602 \times 10^{-19}}$$

$$= 0{\cdot}00001045 \text{ gm./coulomb.}$$

For copper

$$\frac{1{\cdot}660 \times 10^{-24} \times 63{\cdot}57}{2 \times 1{\cdot}602 \times 10^{-19}}$$

$$= 0003294$$

For silver

$$\frac{1{\cdot}660 \times 10^{-24} \times 107{\cdot}88}{1 \times 1{\cdot}602 \times 10^{-19}}$$

$$= 001118$$

The experimental values for these quantities are 0·00001045, 0·0003295, and 0·0011183 respectively.

A convenient conception is that of the Faraday. This constant is the amount of electric charge which will liberate

one gramme-molecule of any substance. For hydrogen it is thus

$$\frac{1\cdot00814}{0\cdot00001045}.$$

By referring back to the equation for the E.C.E. of hydrogen we see that this is

$$1\cdot00814 \Big/ \frac{1\cdot660 \times 10^{-24} \times 1\cdot00814}{1 \times 1\cdot602 \times 10^{-19}},$$

or more simply

$$\frac{1\cdot602 \times 10^{-19}}{1\cdot660 \times 10^{-24}}$$

or $\qquad$ 96490 coulombs.

Faraday's constant must tie up with Avogadro's constant, the number of molecules in one gramme-molecule, for clearly

$$\text{Avogadro's constant} = \frac{\text{Faraday's constant}}{\text{charge on one ion}}$$

$$= \frac{96488}{1\cdot602 \times 10^{-19}}$$

$$= 6\cdot023 \times 10^{23} \text{ per mol.}$$

Evidence for the Ionic Theory

The evidence for the ionic theory may be summarized as follows:

(1) It fits both Faraday's laws of electrolysis.

(2) If a potential difference is applied across electrodes in an electrolyte, the current begins to flow *at once*, which suggests that the molecules are already dissociated with ions. If work were required to separate them, there would be a delay while the current was doing the work. On the other hand, the problem of accounting quite satisfactorily for this dissociation is a complex one.

(3) Bragg has shown experimentally that the crystals of some salts are probably built up of ions rather than of molecules.

(4) The depression of the freezing-point of a solution gives an estimate of the number of separate particles in the solution.

The number found is larger than it would be if all the molecules of the solute were whole; so it seems likely on this ground that some or all of them are split up into ions.

(5) When various secondary effects are known and allowed for, it is found that, on the whole, electrolytes obey Ohm's Law. This suggests that the carriers of electricity, the ions, are all there ready to carry it, and that their total number is not seriously affected by the size of the current flowing, provided that the current is not exceedingly large.

The Copper Voltameter

The Copper Voltameter provides the simplest experimental method of illustrating Faraday's laws.

It consists simply of two copper plates in copper sulphate solution. Pure sheet copper should be used. The electrolyte should contain 25 gm. of copper per litre, the copper being in the form of $KCu(CN)_2$ or $K_3Cu(CN)_4$, with 20% of excess KCN.

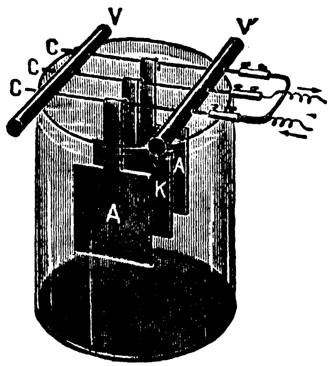


FIG. 86.

Sometimes the anode is in the form of two plates connected together, as in Fig. 86, in order to get equal deposition on both sides of the cathode. The potential difference between the electrodes should be between 3 and 4 volts, and the current should not exceed 0·003 ampere per square centimetre if the copper is to be firmly enough deposited for copper-plating. For about 3% accuracy currents about four or five times as large may be used.

The total charge passing may be found easily in two ways. The current may be kept accurately constant if there is an ammeter or galvanometer, and a finely adjustable rheostat in the circuit. The observer must watch all the time and make continual slight adjustments with the rheostat. Alternatively the current-reading may be watched all the time and recorded as often as possible (say every half-minute).

When enough time has elapsed for a reasonably-measurable mass to have been deposited, a graph is plotted of current against time. The area of the rectangle formed by one horizontal unit of time and one vertical unit of current in amperes represents the passage of 1 coulomb. The total area under the graph (found by planimeter or laborious summation of small areas) gives the number of coulombs which have passed.

The Silver Voltameter

It is more difficult to carry out this process successfully with silver.

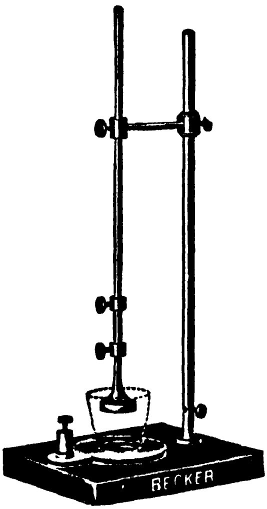


FIG. 87.

The solution contains between 12 and 15 gm. of silver per litre, in the form of $KAg(CN)_2$, with some excess KCN. A little $Ba(CN)_2$ should be added, because atmospheric CO_2 is apt to form K_2CO_3, and the $Ba(CN)_2$ precipitates the carbonate as insoluble $BaCO_3$. The deposit is dull, but can be polished; or the addition of a little carbon disulphide makes the deposit bright. The applied P.D. should be about 1 volt, and the current-density 0·0015 ampere per square centimetre of plate. This method will probably give a value of the electrochemical equivalent about 1% low.

The anode should be a pure silver knob, the cathode a silver or platinum bowl. The apparatus is arranged as in Fig. 87.

The International ampere is defined as the unvarying current which will deposit 0·00111800 gm. of silver per sec. from silver nitrate. This and the International ohm, the resistance of a column of mercury of uniform cross-section, length 106·300 cm. and mass 14·4521 gm. at 0° C., to an unvarying current, are the two fundamental practical standards, the other practical standards being deduced from them. Thus the International volt is the potential difference across an International ohm carrying a current of an International ampere.

The Water Voltameter

Very dilute sulphuric acid can be decomposed by means of the apparatus shown in Fig. 88. Electrodes of thin sheet platinum are fixed in vertical tubes, with stop-cocks at the top of them. These tubes are usually graduated in c.c. from the top. Between 3 and 4 volts are required to drive the current because a back E.M.F. is set up by a sort of cell action which will be explained in the next paragraph. The $H^{\cdot}$ ions go to the cathode, and the SO_4'' ions to the anode. Discharge of SO_4 leads to the secondary equation $2SO_4 + 2H_2O \rightarrow 2H_2SO_4 + O_2$. Since platinum cannot be attacked, the O_2 appears as gas, *two* H_2 molecules (equivalent to $2SO_4$) being formed simultaneously at the cathode.

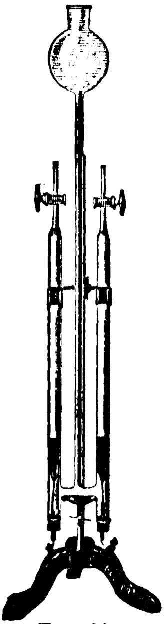


FIG. 88.

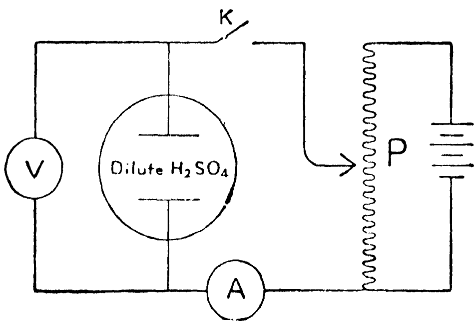


FIG. 89.

Thus, by Avogadro's hypothesis, hydrogen is evolved twice as fast by volume as oxygen. This can be observed in the graduated tubes.

Back E.M.F. in Voltameters

A water-voltameter may be connected up in a circuit as shown in the diagram (Fig. 89), in which K is a key, A an ammeter, V a voltmeter, and P a potentiometer arrangement, consisting of a resistance, across a 6-volt accumulator, with a

sliding contact. This enables any potential difference from zero to 6 volts to be obtained.

With this apparatus results may appear which seem, at first, not to agree with the ionic theory.

If the key K is pressed, and the potential difference across the voltameter is gradually raised from zero, it will be found that little current (and that inconstant and transient) flows through the ammeter until the potential difference is about two volts. After that, the current increases steadily with the potential difference. It would appear, at first, that energy was being used to break up the water-molecules.

If, however, after a current has been flowing, the key K is suddenly opened, so that the voltameter is in series with the voltmeter only, it will be observed that the voltmeter reading falls off gradually, instead of going to zero at once. This shows that the voltameter was temporarily behaving like a cell itself, and producing an electromotive force which was opposing the applied potential difference. The absence of current at first was thus due to a back electromotive force of about two volts, caused by the oxygen on the anode, and the hydrogen on the cathode. These form a cell in which the oxygen is the positive plate, and the hydrogen the negative plate. The first little current which flows, when a potential difference less than two volts is applied across the electrodes, establishes the layers of gas. When they are formed, the current drops to nearly zero, only flowing fast enough to maintain the coatings of gas on the plates against gradual solution in the water.

Secondary Actions

(1) The liberated ion may combine with the electrodes. In the electrolysis of copper sulphate with copper electrodes, the SO_4'' ion on discharge at the anode attacks the copper, forming copper sulphate, which goes into solution. Thus the concentration of the solution as a whole stays constant, because copper out of it is continually being deposited on the cathode. As time goes on the concentration tends to become greater near the anode, and less near the cathode.

(2) The liberated ion may react with the electrolyte.

If caustic soda is electrolysed, sodium is liberated at the

cathode,[1] and OH at the anode. The sodium immediately acts on the water, so that the final liberated product is hydrogen.

$$2Na + 2H_2O = 2NaOH + H_2.$$

While, at the anode, oxygen is finally liberated, for

$$4OH = 2H_2O + O_2.$$

Similarly if a sulphate is electrolysed with a platinum anode, the SO_4'' ion is immediately decomposed, forming sulphuric acid and oxygen, which is the final product.

$$2SO_4 + 2H_2O = 2H_2SO_4 + O_2.$$

It is worth noting that in all cases the final products at the electrodes are—

Metal or hydrogen at cathode.
A non-metal such as oxygen or chlorine at anode.
The following table gives examples :

Electrolyte.	Cathode Product.	Anode Product.
Sulphuric Acid (H_2SO_4)	Hydrogen	Oxygen
Copper Sulphate ($CuSO_4$)	Copper	Oxygen
Nickel Sulphate ($NiSO_4$)	Nickel	Oxygen
Hydrochloric Acid (HCl)	Hydrogen	Chlorine
Copper Chloride ($CuCl_2$)	Copper	Chlorine
Nickel Chloride ($NiCl_2$)	Nickel	Chlorine
Silver Nitrate ($AgNO_3$)	Silver	Oxygen

Practical Applications

The two principal practical applications of electrolysis, both very important, are electro-plating and the production of certain chemical substances in a pure state.

Electro-plating is done by the obvious method of making the substance to be plated the cathode in a cell in which the electrolyte is a salt of the metal to be deposited as plating.

Caustic soda is produced by the electrolysis of brine, pure sodium by the electrolysis of molten caustic soda, and chemically pure copper from crude copper by the ordinary copper

[1] The ordinary view. But see footnote p. 176.

voltameter method. Metals such as aluminium and calcium are difficult to obtain pure by any other method.

The aluminium process is interesting.

Alumina (Al_2O_3) will not dissolve in water, but dissolves in molten cryolite (Na_3AlF_6), which is an electrolyte. A current, so large that the heat generated keeps the cryolite molten, is used for the electrolysis, and the aluminium is liberated at a carbon cathode. The anode is also made of carbon.

In commercial electrolysis, other than direct deposition of a metal from its own salt, secondary actions might cause trouble, if they were allowed to occur; and it is for this reason that platinum and carbon are so often used for electrodes. They are generally immune from secondary action, though carbon has an annoying habit of absorbing gases.[1]

[1] But a very useful habit if one wants a good vacuum.

When a glass vessel is being exhausted by a pump, there comes a time when the pump has done all it can. A small tube attached to the apparatus, containing carbon granules, is immersed in liquid air, which lowers the temperature of the carbon and enormously increases its power to absorb gases. Much of the residual gas is then absorbed, and the tube can be sealed and cut off. This process leaves a much better vacuum than the pump could give unaided.

CHAPTER XII

CELLS

The "Simple" Cell—Local Action and Amalgamation—Polarization and Depolarization—The Daniell Cell—The Leclanché Cell—The Weston and Clark Standard Cells—Miscellaneous Primary Cells—Secondary Cells and the Lead Accumulator—The Iron-Nickel Accumulator—The Theory of Cells and Contact-potential Difference—Amalgamation and Over-voltage.

The "Simple" Cell

IF strips of two dissimilar metals are put in an electrolyte so that they do not touch each other they very rapidly acquire a potential difference. Volta discovered this fact, and used it to make his voltaic pile of alternate strips of copper and zinc separated by wet flannel. The wet flannel functioned because impure water is a weak electrolyte.

The difference of potential of copper and zinc, separated by weak acid, is of the order of 1 volt, and is quite independent of the dimensions of the plates. Large and fairly permanent differences of potential, provided that no current is allowed to flow, are thus easy to obtain by this method.

The underlying theory of this action turns out to be in the province of the physical chemists rather than the physicists; but a brief outline of it will be given at the end of this chapter. The general account will be confined to observable facts and such simple parts of the theory as are directly useful in everyday problems.

Let us consider a cell of copper and zinc in dilute sulphuric acid. The following short series of experiments illustrates the main points which must be grasped about it.

(1) Put a strip of commercial zinc in dilute sulphuric acid. It will dissolve slowly, giving off bubbles all the time.

(2) Take it out again and amalgamate its surface with mercury.

[Put a little mercury on each side of the wet zinc, and try

to rub it off again. You will not succeed, but the whole surface of the zinc will be covered with mercury, which sinks into the zinc in some way without combining with it chemically. The surface thus becomes a sort of mercury-zinc alloy.]

Put the amalgamated zinc back in the acid; you will find that it no longer dissolves, and that it gives off no bubbles.

(3) Put some strip copper in the acid with the zinc. Make the copper and zinc touch outside the acid. Bubbles will immediately begin to appear from the copper addition, and, if you look carefully, you will see that some very little bubbles are clinging to the surface of the copper.

(4) Take the copper out and rub it. Connect the copper to the zinc by copper wires through a low-resistance tangent galvanometer. Put the copper back in the acid, and observe the deflection of the galvanometer.

(5) Watch the deflection for some time. It should decrease. When it has done so, take the copper out, rub it down again with your finger, and put it back. The galvanometer deflection should return to its former value, but should gradually decrease again.

(6) Put some potassium or sodium dichromate crystals into the acid when the galvanometer deflection has fallen off markedly. The galvanometer deflection should jump up and stay up.

Local Action and Amalgamation

In part (3) of the above series of experiments, it was noticed that no bubbles are evolved unless the zinc and copper are in electrical contact outside the acid; that is, unless a current flows. The evolution of bubbles shows that chemical action is going on. This action, therefore, requires a current for its continuance.

In parts (1) and (2) it was found that some action of this kind happens if impure zinc is put in the acid *alone*, but that if its surface is amalgamated with mercury the action stops. The action in impure zinc is called local action, and is due to the formation of innumerable small cells between the zinc and its impurities. Amalgamation covers up the impurities, and prevents local action. [It is difficult to see why this should happen; for mercury ought apparently to form a first-class

impurity and make a cell with the zinc. But it does not. The problem will be considered at the end of the chapter.]

Polarization and Depolarization

The rapid decrease of the current in part (5) is said to be due to polarization. This process consists in the attachment of hydrogen bubbles to the copper plate. These cause the current to diminish for two reasons. They form with the copper an interior cell whose E.M.F. opposes the E.M.F. of the cell as a whole (as when back E.M.F. is produced in voltameters). They also cover the surface of the cell, reduce the area available for the passage of current, and thus increase the internal resistance.

The overcoming of polarization is the chief problem in cells which are to be of commercial value. Theoretically it may be done in two ways: by the continuous brushing away of the hydrogen bubbles by mechanical means, and by their removal through chemical combination with some constituent of the electrolyte. The first method is impracticable, but the second may be used in many forms, of which that of part (6) of the experiment is one. In this instance, the hydrogen is oxidized to form water by the dichromate. We shall now consider in detail the methods of depolarization in various types of cell.

The Daniell Cell

The Daniell cell consists of amalgamated zinc, dilute zinc sulphate (or sometimes sulphuric acid) contained in a porous pot, which stands in copper sulphate with a large copper plate immersed in it. Fig. 90 shows how it is put together.

The internal processes of the cell are as follows:

(1) The zinc ion, $Zn^{\cdot\cdot}$, goes into the zinc sulphate solution, which contains $Zn^{\cdot\cdot}$ and SO_4'' ions. The ion from the zinc may remain an ion, or may combine loosely with the SO_4.

(2) The $Zn^{\cdot\cdot}$ ion goes through the porous pot (which delays mixing of the liquids, but allows ions to pass) into the copper sulphate, where, again, it may or may not combine with the SO_4'', but in any case it sets up an electric field which drives the copper ions toward the copper plate.

(3) The copper ions are deposited on the copper plate and give up their charge to it.

If the first electrolyte is H_2SO_4, the procedure is the same, except that $H^{\cdot}$ ions pass through the porous pot in (2).

The Daniell cell has great advantages and disadvantages.

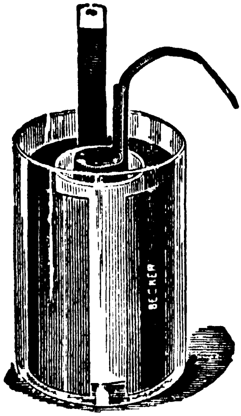


FIG. 90.

The advantages are :

(1) The E.M.F. is so nearly constant at 1·09 volts that it may be used as a rough standard of E.M.F.

(2) It will give a large current—sometimes as much as 3 amps.—for at least an hour without polarizing and with practically no expense.

The disadvantages are :

(1) It cannot be left standing, because in time diffusion takes place through the porous pot in both directions, and the amalgamation tends to wear off the zinc.

(2) It is troublesome to carry about.

(3) It always has to be thoroughly cleaned immediately after it is used.

(4) When it is first put together, it takes a long time, sometimes half an hour, to settle down and give its final E.M.F. and resistance.

The Leclanché Cell

In the Leclanché, the negative pole is again zinc. It is immersed in ammonium chloride (NH_4Cl) which is outside a porous pot.

Inside the porous pot is solid manganese dioxide (MnO_2), and granulated carbon, packed round a carbon rod (Fig. 91).

The $Zn^{\cdot\cdot}$ leaves the zinc with a negative charge as before, and combines with $2Cl'$ to form $ZnCl_2$.

The $NH_4^{\cdot}$ penetrates the porous pot and gives up its charge to the carbon rod. It is unstable when uncharged, and breaks up as follows :

$$2NH_4 = 2NH_3 + H_2.$$

The H_2 reacts with the MnO_2 so that

$$H_2 + 2MnO_2 = Mn_2O_3 + H_2O,$$

thus preventing polarization.

This oxidizing process is a slow one; and so the cell polarizes if it is made to give a heavy current, but it recovers after a rest.

The ordinary dry cell used in a flash-lamp is usually a Leclanché, in which the ammonium chloride is in the form of a paste.

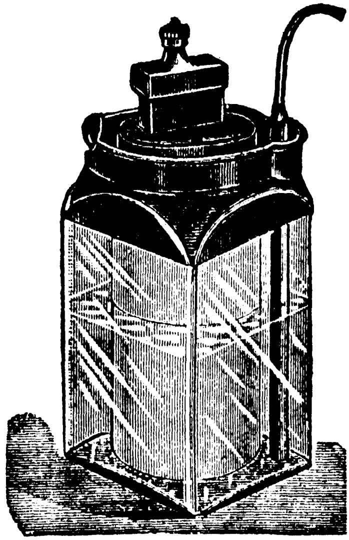


FIG. 91.

The E.M.F. of a new Leclanché is 1·6 volts, but in practice its value usually seems to be below this—probably about 1·43 volts.

The disadvantage of the Leclanché is the obvious one, that it cannot give a heavy current without polarizing.

Its advantage is that it hardly deteriorates at all with age. A battery of Leclanché cells may, in fact, drive the electric bell system of a house for years, with no more attention than the occasional addition of water to the ammonium chloride to replace losses by evaporation.

The Weston and Clark Standard Cells

The better of these is the *Weston* cell. It is usually in the form of an H-tube, and the arrangement of the cell is as shown in Fig. 92.

PP are platinum wires sealed into the glass to make contact with the poles.

M is very pure mercury, the positive pole.

D is a paste of mercurous sulphate, the depolarizer, Hg_2SO_4.

CC are cadmium sulphate crystals ($CdSO_4$) which keep the cadmium sulphate solution (SS) saturated.

A is an amalgam of cadmium and mercury, which forms the negative pole.

The E.M.F. of the Weston cell is 1·0183 volts at 20° C.

The *Clark* cell used to be used as a standard before the Weston cell was invented. It is not so good, because its E.M.F. varies more with temperature.

It is the same as the Weston cell, except that the negative pole is a piece of pure zinc, S is zinc sulphate solution, and C zinc sulphate paste. Its E.M.F. is 1·424 volts at 20° C.

The E.M.F. of the Weston cell varies with temperature, decreasing with rise of temperature above 20° C., and increasing with fall below 20° C.

If the temperature is $(20 + x)°$ C., then the correction is nearly $-(x^2 + 41x) \times 10^{-6}$ volt.

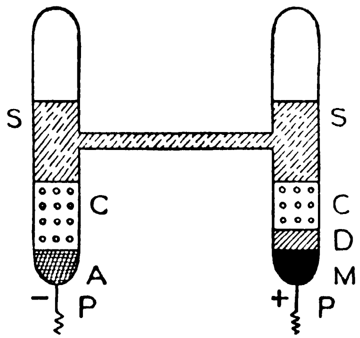


FIG. 92

Both the x^2 and the $41x$ make the E.M.F. diminish if the temperature rises above 20° C., but if the temperature falls below 20° C., so that x is negative, the x^2 term still makes the E.M.F. decrease, but the $41x$ term, which is always larger in ordinary laboratory work, makes the E.M.F. increase.

Thus if the E.M.F. is 1·0183 volts at 20° C., it is at 30° C. [1·0183 − ·00010 − 0·00041] or 1·0178 volts, and at 10° C. it is [1·0183 − 0·00010 + 0·00041] or 1·0180 volts.

The E.M.F. of the Clark cell diminishes by about 0·00172 volt per 1° C. rise of temperature. Its temperature correction is thus, for 1° C., about 43 times as large as that of the Weston.

Miscellaneous Primary Cells

The Chromic Acid cell has a positive pole of carbon and a negative pole of zinc. The carbon is generally double, so that a plate of it is on each side of the zinc. The electrolyte is dilute $H_2SO_4 + CrO_3$ (chromium trioxide).

The main reaction when the zinc dissolves is

$$Zn + H_2SO_4 = ZnSO_4 + H_2.$$

The liberated hydrogen reacts with the chromic acid, which thus acts as depolarizer, as follows:

$$3H_2 + 2CrO_3 + 3H_2SO_4 = 6H_2O + Cr_2(SO_4)_3$$

forming water and chromium sulphate.

The red solution of $H_2SO_4 + CrO_3$ gradually becomes greenish because the $Cr_2(SO_4)_3$ remains in solution. The E.M.F. is about 2·1 volts.

Potassium dichromate may be used instead, but crystals of chrome alum $\{K_2SO_4, Cr_2(SO_4)_3, 24H_2O\}$ are precipitated when the cell is used. Hence it has to be emptied and cleaned more often than for chromium trioxide.

Bunsen's cell consists of zinc-negative and carbon-positive poles like the chromic acid cell, but depolarization is done by the porous pot method.

The carbon pole stands in nitric acid in a porous pot, which in turn stands in dilute H_2SO_4, which also contains the zinc.

The E.M.F. is about 2·0 volts, and the internal resistance is low, but the cell in action gives off the horrible-smelling brown fumes of nitrogen peroxide.

The Grove's cell is the same in every way as the Bunsen cell, except that platinum replaces the carbon as a positive pole.

Secondary Cells and the Lead Accumulator

When the elements of the primary cells, so far considered, are exhausted, they must be renewed. They cannot be restored to their former state by the passage of a current in the opposite direction.

Some cells, however, can be so reversed, and they are called secondary cells, since they do not give up their own energy directly, but receive electrical energy, store it as chemical energy, and release it again as electrical energy. These secondary cells, or accumulators, are regarded sometimes as stores of electrical energy; but this view is quite mistaken, for their energy is entirely chemical, though it is transformed to electrical energy as it leaves.

Strictly speaking, in the physical sense, a lead accumulator is not a reversible cell. The charge equations are not the reverse of the discharge equations, and the E.M.F. on charge

is not the same as the E.M.F. on discharge. An accumulator is only reversible in a colloquial sense. It is more accurately described as a polarizable cell which can hold its polarization for a long time. A Daniell cell is more nearly reversible, and is usually called reversible. But it is not, for the hydrogen lost on discharge cannot be regained on charge.

The positive and negative poles of an ordinary accumulator are connected to strong lead plates. On the positive plate is a hard brittle substance, chocolate-brown in colour, called lead peroxide (PbO_2). This is known as the "positive active material."

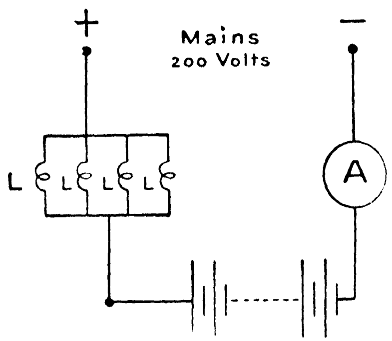


FIG. 93.

On the negative plate is spongy lead; that is pure lead in a very porous condition. It is called the "negative active material." Dilute sulphuric acid is the electrolyte. In small accumulators, the specific gravity of this acid is 1·250 when the cell is fully charged, and about 1·180 when it is discharged.

The cell is charged by having a current sent through it from positive to negative, at a rate depending on the size of the plates, and invariably stated on the maker's label. It does no harm to charge below this rate, though, of course, charging takes longer, but it does a lot of harm to charge too fast, for the active material tends to get blown off the plates by the too rapid formation of gas. Of course the E.M.F. of the accumulator opposes the charging E.M.F.

The usual (and wasteful) method for small accumulators is to charge from direct-current mains through carbon-filament lamps of resistance 200 ohms. Such lamps pass 1 amp. from 200-volt mains.

The usual circuit is shown in Fig. 93. LL are lamps, and A is an ammeter. If each of these lamps had 200 ohms resistance, the current, *before the accumulator was put in,* would be 4 amps.

If the voltage of the accumulator was 20 volts, then, after it was put in, the E.M.F. in the circuit would be (200 − 20) or 180 volts. So the current would be 3·6 amps., instead of 4·0 amps.

A typical charging curve of an accumulator is shown in Fig. 94.

Though, while charging is going on, the E.M.F. may rise as high as 2·65 volts, it drops to 2·08 volts very soon after the charging is stopped.

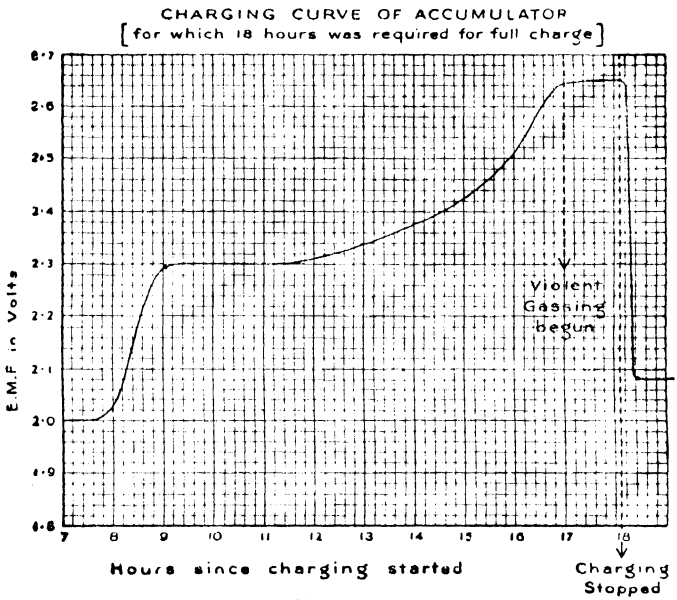


FIG. 94.

In a discharged or uncharged accumulator there is a hard deposit of lead sulphate on top of the active material.

The energy of the charging current is used to decompose this deposit. This drives the SO_4 back into the solution, thus increasing the density, and leaving pure lead peroxide on the positive plate and spongy lead on the negative plate. When all the sulphate is removed the energy of the current

occupies itself in decomposing the water of the solution and this causes the gassing.

The discharge equations are:

Positive plate $PbO_2 + H_2 + H_2SO_4 = PbSO_4 + 2H_2O$.
Negative plate $Pb + SO_4 = PbSO_4$.

The charge equations are:

Positive plate $PbSO_4 + SO_4 + 2H_2O = PbO_2 + 2H_2SO_4$.
Negative plate $PbSO_4 + H_2 = Pb + H_2SO_4$.

Reference to authorities will show that there are many versions of the equations of a lead accumulator. The equations given above are well represented in these authorities.

When gassing starts, therefore, it means that the accumulator is charged. If charging is continued long after gassing has started, lead peroxide is blown off the positive plate by the gas forming under it. It falls towards the bottom of the case, and some of it on the way down gets deposited on the negative plate, forming " trees " which may in time grow across and cause a short-circuit.

The best way of estimating the effective " age " of an accumulator is by the amount of deposit on the bottom of the case. An overcharged accumulator grows old before its time.

When the terminals of an accumulator are connected through a resistance, a current flows, receiving its energy from the reconversion of lead peroxide and spongy lead into lead sulphate. The specific gravity of the acid and the E.M.F. of the cell both fall.

If the cell is made to give too big a current for the size of its plates, the plates may buckle, through the production of too much heat, or through the change of all the active material on the surface to lead sulphate before the lower deposits are affected.

If the discharge is continued too long, basic lead sulphate is formed. Even in mild cases this may take a long time to remove, and in bad cases it is impossible to remove it, and the cell is ruined.

Fig. 95 shows a typical discharge curve of a cell giving its correct current at such a rate that it would be completely discharged in 10½ hours.

It should be noticed that the E.M.F. falls from 2·08 volts to 2·0 volts almost immediately, and that, when it has reached 1·9 volts, it has only an hour or so more to go; in fact, for more than three-quarters of its time it is between 2·0 volts and 1·95 volts.

It is well to observe that it is impossible to distinguish with a voltmeter between a fully charged battery and one

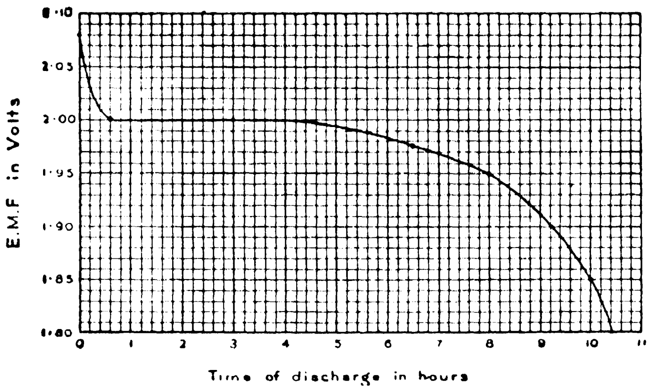


FIG. 95

which is half discharged (a fact which has caused much profit to dishonest garage proprietors who trade on the ignorance of their customers).

The specific gravity of the acid, on the other hand, falls almost in proportion to the time of discharge. So, if you want to find out whether your battery has been properly charged, use a hydrometer rather than a voltmeter. Even this is not safe, for it is possible to add acid and send up the specific gravity without completing the charge.

On the whole, it is well worth while to stick as accurately as you possibly can to the maker's instructions for an accumulator. If you do not, you will soon have to buy a new one, for accumulators do not forgive mistakes.

One of the chief merits of the accumulator is that its

internal resistance is very small indeed—probably of the order of one-thousandth of an ohm if it is in good condition. [It may rise a lot if the accumulator is badly sulphated.] Its resistance is thus less than one two-hundredth of that of a good Daniell or Leclanché cell.

The *capacity* of an accumulator is the quantity of charge it will pass before it is so much discharged as to need re-charging. It is measured in ampere-hours, the ampere-hour being the amount of charge which passes when 1 amp. flows for 1 hour. One ampere-hour is thus 3,600 coulombs.

A 60 ampere-hour (actual) accumulator should thus be able to give 2 amps. for 30 hours, or 3 amps. for 20 hours, and so on until the maximum discharge-rate is exceeded. If the accumulator were discharged at more than its maximum discharge rate it would give fewer ampere-hours.

The ampere-hour efficiency, or the ratio of the charge an accumulator will give out to the charge it must receive to bring it back to a fully-charged condition is about 90 per cent. for a good make, and it may be more.

The watt-hour efficiency, or ratio of the energy the accumulator will give out to that it must receive, is only about 75 per cent.

The difference between the two efficiencies is due to the fact that the E.M.F. of the accumulator on charge for a given current is on the average about 20 per cent. more than it is on discharge.

Only distilled water should be used in the electrolyte both for the original filling and for later replacement or " topping-up," which is necessary because water is lost by evaporation.

The acid does not escape except by spilling or leakage, and therefore does not in general need replacement. In general it is better, if some acid has been spilt, to empty out the whole lot of the electrolyte, and refill the cell with new acid and water. In any case this should be done periodically.

It is most important not to exceed the maker's rates of charge and discharge.

The moment when an accumulator is completely discharged is easy to detect with a voltmeter, although the voltmeter does not distinguish between one-quarter discharge and three-quarters discharge.

For any particular discharge-rate there is a safe minimum voltage. [Discharge-rate is here used not in the sense of current, but of the number of hours taken to discharge the accumulator completely. Thus the 8-hour rate is the rate at which the accumulator is completely discharged in 8 hours.]

The greater the discharge-rate (a 2-hour discharge-rate being regarded as a greater rate than a 4-hour) the lower the voltage may be taken, and the greater, therefore, the ampere-hour efficiency of the accumulator.

A minimum voltage of 1·83 volts per cell at the 10-hour rate to 1·75 at the 1-hour rate is approximately right. Thus if the voltage of an accumulator discharging at the 10-hour rate be observed *when the current is flowing*, it may be regarded as discharged when 1·83 volts is observed.

The Iron-Nickel Accumulator

This interesting type of accumulator is sometimes called the "Edison" cell. Its positive plate has a nickel hydroxide (possibly, but not certainly, $Ni(OH)_4$) as active material, and iron is the negative active material.

The electrolyte is a 21 per cent. solution of KOH in water, to which is added about 50 grammes per litre of lithium hydrate. The function of this is not known, but it increases the capacity. On discharge OH′ ions from the KOH go to the negative plate and react with iron according to the reaction

$$Fe + 2OH' \longrightarrow Fe(OH)_2,$$

and $K^{\cdot}$ ions go to the positive plate, with the reaction

$$2K^{\cdot} + Ni(OH)_4 \rightarrow Ni(OH)_2 + 2KOH,$$

so that the OH lost at the negative plate reappears at the positive plate, and the constitution, and hence the specific gravity, of the electrolyte remains unchanged.

All that happens is that OH is transferred from one plate to the other. On charge the reverse process happens.

Thus the whole equation of the cell is

$$Ni(OH)_4 + KOH + Fe \rightleftarrows Ni(OH)_2 + KOH + Fe(OH)_2.$$

The E.M.F. of the cell is about 1·4 volts, but it drops uniformly to about 1·0 volt as the cell discharges. At the end of a charge the P.D. is about 1·8 volts.

The ampere-hour efficiency is only about 75 per cent., and the watt-hour efficiency about 60 per cent. The watt-hours per cubic foot of space occupied are less than for a lead accumulator, and the internal resistance is higher.

On the other hand, the watt-hours per pound weight are higher than for the lead cell, and the accumulator is far more robust, both electrically and mechanically. It can be short-circuited or discharged at a high rate, or left discharged, without damage, and it can be charged at very high rates. It is difficult to break, since the negative plates are iron, and the positive plates are nickel welded on to steel. It is thus useful for rough work such as haulage, but comparatively useless for large stationary storage work such as house-lighting, owing to its large volume and wide variations of E.M.F.

The Theory of Cells and Contact-potential Difference [1]

Two elements in contact may show a considerable potential difference; 0·75 volt for zinc and copper. This varies with temperature; so in a circuit composed of two metals an E.M.F. acts when the junctions are at different temperature, see p. 169.

The element which is the more positive of the two is said to be electropositive to the other, and the negative metal is said to be electronegative to the first.

It is possible to arrange many elements in a particular order in which (with occasional small variations in special circumstances) the element higher in the list is electropositive to any lower element.

Electropositive end.

Cæsium	Nickel	Rhodium
Rubidium	Thallium	Platinum
Potassium	Indium	Osmium
Sodium	Lead	Silicon
Lithium	Cadmium	Carbon

[1] The theory outlined on pp. 210–211 does not directly require the discussions of pp. 206–209. But this seems a good place to discuss contact-potential difference. For compare experiments 2 and 3, p. 214.

Barium	Tin	Boron
Strontium	Bismuth	Nitrogen
Calcium	Copper	Arsenic
Magnesium	Hydrogen	Selenium
Aluminium	Mercury	Phosphorus
Chromium	Silver	Sulphur
Manganese	Antimony	Iodine
Zinc	Tellurium	Bromine
Gallium	Palladium	Chlorine
Iron	Gold	Oxygen
Cobalt	Iridium	Fluorine
		Electronegative end.

If plates of two of these elements are in an electrolyte, the more electronegative element forms the positive pole of the cell, and the more electropositive element forms the negative pole.

Thus zinc and copper form a cell with copper positive and zinc negative, but a cell of higher E.M.F. is obtained by using carbon or platinum in place of copper, since the separation on the list between the two elements of the cell is then greater.

In general, when a cell supplies a current, the electropositive element (the negative pole) is attacked, and hydrogen or an equivalent metal is deposited on the electronegative element (the positive pole).

Let us now consider the nature of contact potential difference. It involves problems which are as yet unsolved.

There are wide variations in the measurements of its value by independent observers. The zinc–copper contact P.D., for example, has determinations varying from 0·71 volt to 0·85 volt. On the other hand, there is general agreement about temperature variations as manifested in the thermoelectric circuit, whose total E.M.F. depends on the difference between the contact P.D.s of two junctions at different temperatures.

If a single contact P.D. is being measured (with a quadrant electrometer, for example) it is difficult to know whether the P.D. is really between the two metals or their surface films, for a film of the metallic oxide is not easily detected or removed.

Experiments indicate that contact P.D. is independent of the surrounding air-pressure, but that the direction of the P.D. for iron-copper is actually reversed if air is replaced by hydrogen sulphide.

The contact P.D. also diminishes in air if the air is dried, which suggests that the air-moisture behaves as if it were a weak electrolyte; it is also reduced at low temperatures in air. Contact P.D. varies if the surfaces in contact are polished.

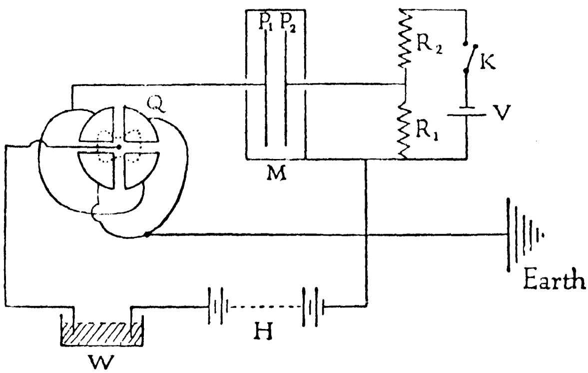


FIG. 96.

Fairly constant values of contact P.D. are obtained if fresh surfaces of the metals are cut *in vacuo*. Such values are—

Aluminium–Copper	+ 0·83	volt.
Zinc–Copper	+ 0·76	,,
Lead–Copper	+ 0·55	,,
Brass–Copper	+ 0·15	,,
Gold–Copper	− 0·11	,,
Silver–Copper	− 0·13	,,

It is found, within the limits of experimental error, that these values are additive. For example, Aluminium-zinc gives (0·83–0·76) or 0·07 volt, and Aluminium-silver gives $\{0{\cdot}83 - (-0{\cdot}13)\} = 0{\cdot}96$ volt.

The measurement of contact P.D. is carried out as in Fig. 96, in which Q is a quadrant electrometer, H a high-tension battery, W a water-resistance, V a 2-volt cell, R_1 and R_2 resistance-boxes, P_1 and P_2 plates of the

different metals whose contact P.D. is to be found, M an earthed metal screen to cut off external fields, and K a key. P_1 is earthed and then insulated. The ratio $\frac{R_1}{R_2}$ is adjusted until the electrometer is unaffected by a small relative movement of the plates P_1 and P_2. The applied potential on P_2 must then cancel out the contact potential. As the applied potential is $\frac{-R_1}{R_1 + R_2} . V$, it follows that the contact-potential of P_2 with respect to P_1 is $\frac{+R_1V}{R_1 + R_2}$.

If, however, the potential of the plate P_2 is altered, by adjusting the resistances R_1 and R_2, until there is no change in deflection when P_2 is moved away from P_1, then the plates will be at the same potential. This will mean that the potential applied to P_2 is equal and opposite to its contact potential with respect to P_1. This applied potential is clearly

$$\frac{R_1V}{R_1 + R_2},$$

which gives the value of the contact potential difference.

Contact potential difference appears as a factor to be considered in photo-electric measurements. If light shines on a metallic surface, electrons are shot out of the surface. The sum of their kinetic energy and the work done in getting them out of the surface are proportional to the frequency of the light. A graph of kinetic energy against frequency gives a straight line which does not go through the origin. The distance, along the energy axis, of the point where the line crosses, represents the work done in getting clear of the surface.

This is the work done in making an electron, whose charge is known, fall through a potential V. If this potential is V_1 and V_2 for two different metals, this $(V_1 - V_2)$ is, within the limits of experimental error, equal to the contact P.D. between the two metals if their surfaces are freshly cut *in vacuo*.

All these results show that though something fairly definite is known about contact P.D.s with freshly-cut surfaces *in vacuo*, our knowledge of them under the conditions of working of cells is far too scanty to form the basis of a complete theory.

A general theory of the behaviour of cells is provided by the theory of Electrolytic Solution Pressure. When a vapour is in contact with the surface of its own liquid, the average rate at which molecules from the vapour cross the boundary surface into the liquid is equal to the average rate at which molecules of the liquid escape from this surface.

There is thus equilibrium between the molecules crossing the surface inwards and those crossing it outwards.

The pressure of the vapour in equilibrium with the liquid is called the *Saturation Vapour Pressure.*

If a solution is inside a porous pot, in which a semi-permeable membrane has been deposited, which allows the molecules of the solvent to pass, but not those of the solute, we may regard the porous pot as a liquid surface of the solute only.

If the porous pot stands in a bath of solvent so that the levels of solution inside and solvent outside are the same, it is obvious that the concentration of solvent molecules is greater outside the pot than inside it.

Therefore molecules of the solvent will pass into the pot more rapidly than those inside pass out, until equilibrium is established with the level of the solution higher than that of the solvent. The absolute pressure of the solution will then be greater than that of the solvent.

Clearly the excess pressure will be due to the molecules of the solute.

This excess pressure is known as the *Osmotic Pressure,* and is of far-reaching importance, particularly in Bio-Chemistry.

The property of a substance which determines its tendency to go into solution is known as the *Solution Pressure.*

As the Saturation Vapour Pressure is the pressure exerted by a vapour in equilibrium with its liquid, so the Solution Pressure of a substance is the Osmotic Pressure of its saturated solution in equilibrium with the solvent. The conception of Solution Pressure may be applied to the passage of substances into the ionic condition when in contact with an electrolyte.

If a metal is dipped into water it has a tendency to dissolve, just as water has a tendency to give off water vapour. A metal can only dissolve in ionic form, and it thus goes into solution in the water in the form of positive ions. The water (now a weak

solution) thus becomes positively charged, and the metal becomes negatively charged (usually, but not invariably).

After a very short time, when the concentration of the metallic ions is exceedingly small, the water becomes positive enough to prevent any more ions from coming into it.[1]

If the metal is dipped into a solution of one of its salts, the situation is altered, since positive metallic ions already exist in the solution. What happens will depend upon the relative values of osmotic pressure of the ions in the solution and the solution pressure of the metal.

If the osmotic pressure is greater, positive ions are driven on to the metal, which becomes positively charged. If the solution pressure is greater, positive ions enter the solution, and the metal is negatively charged.

The base metals, such as zinc and iron, have high solution pressures, and are thus negative with respect to solutions of their salts. The noble metals, such as silver and mercury, have low solution pressures, and are usually positively charged with respect to the solution.

Thus in general two metals dipped into an electrolyte are at different potentials. If they are externally connected by a conductor, a current flows. Its magnitude depends on the potential difference between the electrodes, on the resistance of the electrolyte, and on the external resistance. Lord Kelvin thought that the E.M.F. of the cell could be calculated from the heat of solution of the electrode metals. This idea was a mistake, as is shown in Appendix IV, p. 494. Kelvin's method of calculation, given below as being of historical interest, happens accidentally to give the correct result for the Daniell cell, which neither emits nor absorbs heat when it works.

When 1 coulomb passes through a Daniell cell, 0·00034 gm. of zinc goes into solution, and 0·000329 gm. of copper goes out of solution. The heat of solution of zinc is 1630 calories per gm. For copper it is 881 calories per gm. Thus the energy available

$$= 0{\cdot}00034 \times 1630 - 0{\cdot}000329 \times 881 \text{ calories,}$$
$$= 0{\cdot}264 \text{ calorie,} \quad = 1{\cdot}10 \text{ Joules.}$$

Thus since 1·10 Joules are dissipated when 1 coulomb passes,

[1] But when the liquid originally contains no ions of the electrode metal but does contain foreign ions (as for mercury in dilute sulphuric acid) a more complex type of equilibrium is set up.

the E.M.F should be 1·10 volts, which agrees well with the observed value.

For other cells this method is of no value.

The general problem of the E.M.F. of cells is discussed from the thermodynamical point of view in Appendix IV, p. 493.

We have seen that when a metal and an electrolyte are in contact there is in general an electrified double layer at the surface of contact.

The surface tension of a liquid can be shown to diminish when it is electrified.

Thus when mercury is in contact with an electrolyte its surface tension is reduced. Since, as we have seen, mercury is, in general, positively charged in contact with an electrolyte, its surface tension can be still further reduced by making it still more positive.

If, however, the mercury is given a gradually increasing negative potential with respect to the electrolyte, the surface tension gradually increases, until it reaches a maximum when the applied P.D. is equal and opposite to the contact P.D. When this happens the double electrified layer is destroyed, and the mercury surface is not electrified. If the negative potential of the mercury is increased, the layer is reversed, and the surface tension is diminished again.

This principle is employed in the Capillary, or Lippmann, Electrometer.

Two vertical glass tubes are joined by a slightly sloping capillary tube. Platinum wires are sealed into their lower ends. Dilute sulphuric acid is poured into the apparatus until it fills the capillary.

Mercury is then poured into the vertical tube at the lower end of the capillary until it reaches half-way up the capillary.

A little mercury is poured into the second vertical tube so that the two platinum wires are at the same potential.

The position of the mercury meniscus in the capillary now depends upon the surface tension.

If the mercury meniscus is made more positive, the surface tension is reduced, and the meniscus moves up the capillary.

If the mercury meniscus is made negative, the surface tension

is increased until the potential reaches about 1 volt, after which it is diminished again.

The whole problem may also be considered from quite a different point of view.

The electrified double layer between the mercury and the acid may be regarded as forming a condenser charged to the contact P.D.

Now in any condenser of constant P.D. the electric forces act in such a direction as to increase the capacity. With plates of fixed area, they draw the plates together. With plates at a fixed distance apart, as in this case, they increase the area if this is possible.

The surface tension, on the other hand, acts in such a direction as to reduce the surface area, which may be imagined to extend along the walls of the capillary, since the mercury is pulled out of the capillary by the surface tension.

Thus the effect of the electric forces is to reduce the surface tension. Therefore when the applied P.D. opposes the contact P.D., and the electric force is reduced, the surface tension increases until the electric force is zero. The applied and contact P.D. are then equal and opposite.

The capillary electrometer is easy to make for demonstration purposes. It only gives accurate quantitative results if it is very clean. It is difficult to get the mercury and the capillary tube clean enough.

A number of easy experiments can be performed with very simple apparatus to demonstrate the theory outlined in this section.

1. A beaker of distilled water is placed in a Faraday ice-pail connected to a sensitive electroscope. A sphere of paraffin wax suspended from an insulating thread is dipped into the water and removed. The water is found to have a positive charge. If the sphere is replaced in the ice-pail, the deflection of the electroscope returns to zero because the sphere has a negative charge equal to the positive charge on the water. The separation of the two parts of the electrified double layer causes this. A substance like paraffin wax, which does not remain wet all over, is obviously necessary. This experiment may be regarded as demonstrating how the so-called "frictional" charges are obtained.

2. A zinc plate and a copper plate are fixed together so that they stand vertically about a centimetre apart.

A small blob of solder is suspended by a light metallic filament midway between them.

A high-tension battery is arranged to give a potential difference of 100 volts or more between the filament and the plates. A high resistance must be connected in series with the battery to protect the filament.

If the filament is observed with a microscope, it is seen to be attracted to the copper plate when the filament is positive, and to the zinc plate when the filament is negative. This result shows that there is an electric field from zinc to copper. If the filament is not accurately centred, it may be attracted to the nearer plate in both cases. Accurate centring is the most difficult part of the experiment.

3. The copper plate in the above experiment is replaced by a filter paper fixed to a vertical wooden support. Both the zinc plate and the filter paper are immersed in an electrolyte at their lower ends. The deflection of the filament shows that the zinc is negative to the solution sucked up by the filter paper.

The object of the filter-paper arrangement is to obtain a vertical column of electrolyte.

4. Platinum electrodes dip into the two arms of a U-tube containing distilled water with lycopodium powder in suspension. An electric field is produced between the electrodes by means of a Wimshurst machine. The lycopodium grains move to the anode, since they are negatively charged with respect to the water. The movement of particles in suspension to the anode is known as *cataphoresis*.

5. The lycopodium suspension in the last experiment is removed. Some cotton-wool is put in the bend of the U-tube, and the U-tube is filled with distilled water. The level of the water is found to rise very slightly at the cathode. This happens because the water is positively charged with respect to the solid cotton-wool plug.

6. A vertical flask containing distilled water is arranged with one electrode at the bottom and the other at the top. The two electrodes are connected to a galvanometer. Porcelain beads are dropped into the water, and fall to the lower

electrode. A current flows through the galvanometer in such a direction that the upper electrode is seen to be positive and the lower one negative. Again the solid beads were negatively charged with respect to the electrolyte.

7. Wire netting electrodes are placed above and below a piece of wet peat in contact with it. The water drips out when the lower electrode is the cathode, but not when it is the anode, since the water is positive with respect to the peat.

8. A copper plate is fixed in concentrated $CuSO_4$ in a porous pot. The porous pot is placed in dilute $CuSO_4$. Another copper plate is fixed in the dilute solution. The first plate is found to be slightly positive to the second, since the solution pressure of the copper ions in the concentrated solution is obviously greater than it is in the dilute solution.

This arrangement is called a *Concentration Cell.*

9. Water is allowed to flow from a burette. When the level has fallen enough, the water comes out in drops instead of in a jet.

The water is connected to a Wimshurst machine. It will be found possible to turn the Wimshurst at such a speed that the water begins to flow in a jet again, since the electrification of the surface reduces the surface tension and prevents the formation of drops. If the machine is revolved fast, however, the jet breaks up into drops which fly out almost horizontally, owing to the electrostatic repulsion of the large charges for each other.

10. A drop of mercury, in a watch-glass, is covered with dilute H_2SO_4. It at once lies much flatter because the surface tension is reduced by the appearance of the electrified layer.

A large iron nail is now fixed vertically in a retort-stand so that it just touches the top of the mercury near its edge.

The nail and the mercury form a short-circuited cell, and the electrification of the double layer is reduced. The surface tension therefore increases. The drop contracts and ceases to touch the nail. At once its surface tension is reduced, and it falls back and touches the nail again.

When the right adjustment of the nail is found, the drop continues to pulsate for long periods; sometimes for half-an-

hour or more. A single crystal of bichromate should be dropped into the acid to prevent polarization.

This experiment ("the beating heart") is particularly well worth seeing.

A more detailed account of these experiments, and others like them, will be found in Pohl's *Physical Principles of Electricity and Magnetism*, except for the last experiment, the account of which is more detailed here.

Amalgamation and Over-voltage

Theoretical physical chemistry indicates that hydrogen should be evolved in electrolysis when the applied P.D. exceeds 1·1 volts. This is nearly true with platinum electrodes. With electrodes of other metals the applied P.D. required is higher. The excess of the applied P.D. over the theoretical value is called the over-voltage of the metal. Platinum thus has the lowest over-voltage.

It is now obvious that the current passed by the cell depends very much on the over-voltage of the positive pole (or electro-negative element). Polarization and over-voltage are thus intimately related.

The over-voltage of mercury is particularly high. It is therefore supposed that, when mercury and zinc are amalgamated, the local action which ought to take place between them is stopped by the over-voltage of the mercury. On this theory the amalgamated mercury would prevent local action between the other impurities and the zinc, by covering them up, and have its own local action prevented by its over-voltage.

CHAPTER XIII

ELECTROMAGNETISM AND ELECTROMAGNETIC INDUCTION

Magnetic Field of a Straight Current—Magnetic Field of a Circular Coil—Solenoids—Electromagnets—Forces on Soft Iron in a Magnetic Field—Force on a Wire carrying a Current in a Magnetic Field—Fleming's Left-hand Rule—Barlow's Wheel—Stretched Wire in a Magnetic Field—Couple on Coil carrying Current in a Magnetic Field—Forces between Currents—Induction of Currents—Magnetic Flux through a Circuit—Lenz's Law—The Five Laws relating Currents and Magnetic Fields—Induction between Currents—Starting and Stopping Currents—Self-inductance and Mutual Inductance—Demonstrations of Self and Mutual Induction.

Magnetic Field of a Straight Current

IT is vital, in this part of the subject, to grasp the principles by doing, or at least seeing, as many simple and fundamental experiments as possible.

Such experiments will, therefore, be described in great detail.

An electric current, or a moving electric charge, produces a magnetic field which acts cylindrically round the wire carrying the current, or round the direction of motion of the charge. This fundamental relation, the link between electricity and magnetism which was sought for 200 years unsuccessfully, was found, almost accidentally, by Oersted in 1820.

The direction of the field is related to the direction of the current as shown in Fig. 97.

If a heavy piece of copper wire is fixed vertically by means of a wooden retort-stand, so that it pierces a horizontal piece of cardboard, the lines of force may be shown by sprinkling iron filings on the cardboard and tapping gently. If one does not mind getting a good many filings on the table or the floor, the best way to sprinkle them is to blow them on to the cardboard from a small heap on another piece of cardboard. This method makes it sure that one gets only very

small particles, which show the field best. A current strong enough to show the field is difficult to get. The effect can just be seen with 5 amps., but 10 amps. or 15 amps. is much better. Such large currents can hardly be obtained without

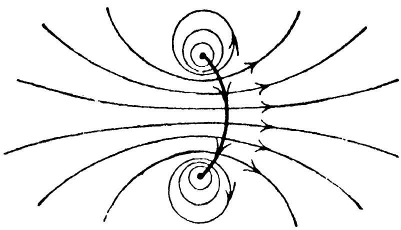


FIG. 97.

D.C. mains. The simplest method of getting such a current without special facilities is to use the terminals of a carbon-arc system, since the series-resistance and the fuses must necessarily be arranged to allow a short-circuit across the terminals by a piece of thick copper wire.

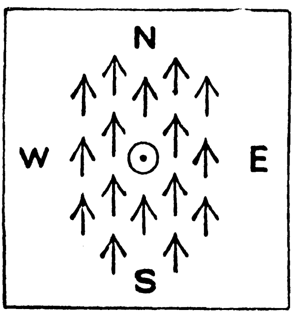


FIG. 98.

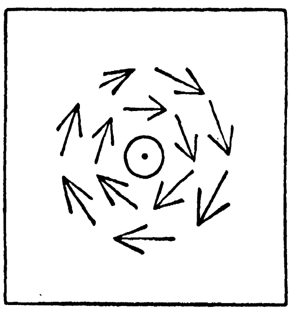


FIG 99.

The direction of the current can be shown by putting a number of small compasses on the cardboard. When the current is off, they all point to magnetic north as in Fig. 98. When the current is on, they point as in Fig. 99, in which the direction of flow of the current (from positive to negative) is vertically downwards through the dot in the centre of the

paper. The direction of the field is then as shown by the arrows as in Fig. 100.

If a right-handed corkscrew be imagined as being screwed into the central dot in Fig. 99 or Fig. 100, the ends of the handle will turn in the direction of the arrows. Thus the relation of rotation of a right-handed cork-screw to translation through the cork is the same as the relation of the magnetic field to the direction of flow of the current causing it.

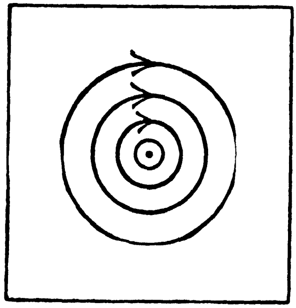


FIG. 100.

Magnetic Field of a Circular Coil

If the wire be wound in a circular coil, and this coil be fixed in a vertical plane piercing a horizontal piece of card-

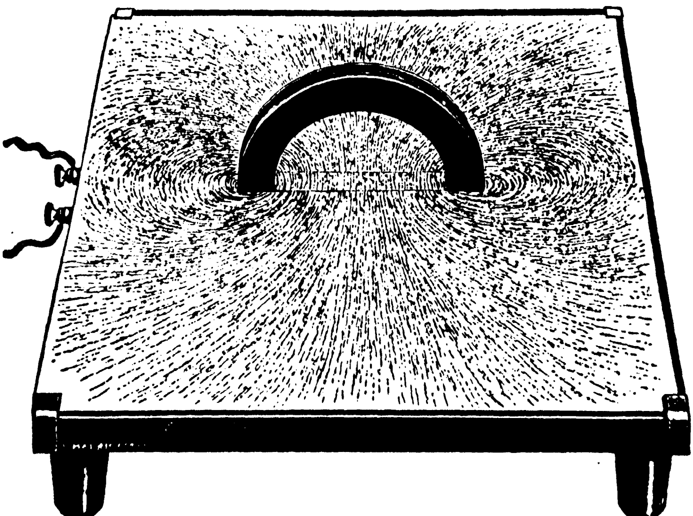


FIG. 101.

board at opposite ends of its diameter, the direction of the field will be as shown in Fig. 97. This can be demonstrated with iron filings as in Fig. 101. The general shape of the

field is obviously what one would expect from the field of a straight wire.

The experiment of Fig. 101 gives one a cross-section of the field; but the field can only be properly observed in three dimensions.

To observe the field satisfactorily, two special coils should be wound to suit the local supply of current. Only one of these is really necessary to show the shape of the field, but two are necessary for later experiments.

The coils described are used in the writer's laboratory, which has 200 volt D.C. mains. Of course they must be designed to suit the local mains.

The two coils are circular, and made of d.c.c. copper wire, gauge 20. Their diametrical cross-section is as shown in

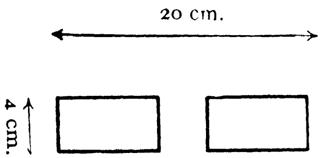


FIG. 102.

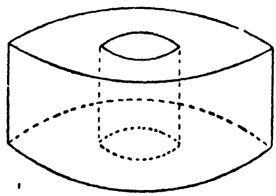


FIG. 103.

Fig. 102, and their general outline as in Fig. 103. The mass of wire used is 10 lb. for the two in the writer's laboratory.

The coils are permanently in series, being connected by a long strong piece of single flex. The two spare ends of the coils are connected to a piece of double flex, at the end of which is a plug for the demonstration-bench circuit.

The heating effect is too great if one of these coils is put straight across the mains by itself, but if the two are put in series the effect is just not too great, provided they are never on for more than five minutes consecutively.

This pair of coils will be referred to in future as the "big" coils.

The big coils can be used to show the distribution of the complete field round a coil. One of them, carrying a current, is brought near a large sheet of drawing-paper on which is a really satisfactory heap of iron filings. As the coil approaches the filings begin to sit up and take notice. Then they leap

up to the coil and drape themselves round it. Finally, all may be picked up, and they are found to show the arrangement of the field most conspicuously.

Many more experiments may be done with the big coils.

If a large flat glass bowl of water of about 18″ diameter, with the water about 4″ deep, be balanced on the coil, an instructive experiment can be done with paper boats.

Small paper boats are built and loaded with iron filings. They are put round the circumference of the bowl before the current is switched on. When the current is switched on they will sail quickly to the centre and there sink vertically, roughly describing the curve of the field.

Iron filings floating on the surface of the water will be found to sink directly the field appears, if the water is shallow enough.

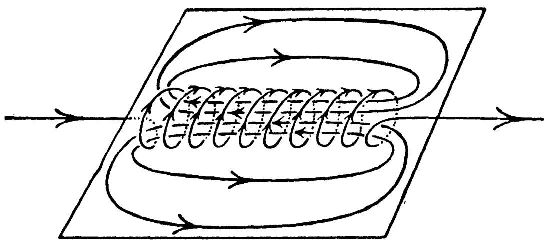


FIG 104.

If a coil is fixed vertically about 18″ above the bench, and a wide glass tube, just not too wide to go through the hole in the middle of the coil, is fixed vertically through this hole with one end on the bench, it will be found that soft iron rods dropped into the top of the tube oscillate furiously about the centre of the coil, and finally stop suspended in mid-tube.

If the coil is held with its axis perpendicular to the horizontal component of the earth's magnetic field, and is "shone" like a torch on a magnetometer some yards away, it will be found to produce a deflection. The extension of its effective field in space is thus realized.

Solenoids

A helical (or spiral) spring of very flat pitch is a solenoid. Its field may be illustrated by winding a coil of stiff copper

wire through holes in a piece of cardboard, so that half the coil is above the cardboard and half below, as in Fig. 104.

If the direction of the current follows the little arrows on the coil, that of the field must clearly follow the big arrows. The direction of the field inside the solenoid is thus along the

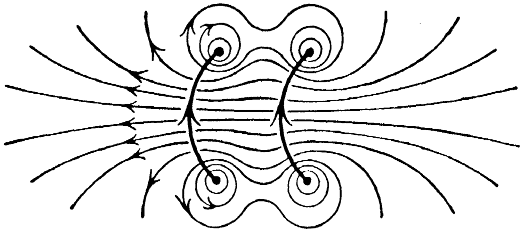


FIG. 105.

axis in such a sense that the corkscrew rule can be applied the other way round. Thus if we consider the motion of translation of the corkscrew inside the coil to be that of the field, then the motion of points in the handle gives the direction of the current in the solenoid.

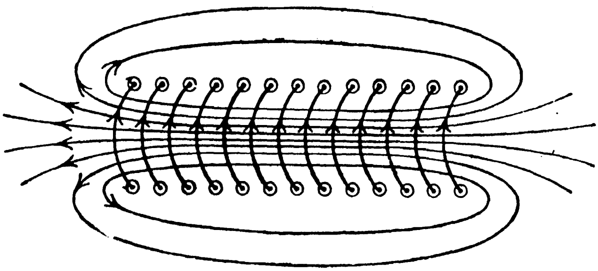


FIG. 106.

If the field due to the solenoid is plotted with iron filings it will follow the general scheme of the diagram, except that interesting eddies will appear wherever the wire pierces the cardboard, as shown in Fig. 105.

The magnetic field of a solenoid has a very close resemblance to the magnetic field of a bar-magnet; and this resemblance is best understood by a direct comparison of the appearances of their fields.

Fig. 106 shows the field of an empty solenoid, and Fig. 107

the field of a solenoid containing a piece of steel which has been magnetized to form a bar-magnet. The general appearance of

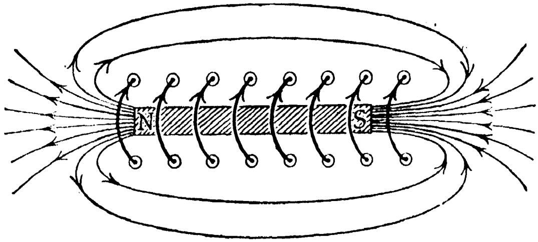


FIG. 107.

this field will not be found to change if the current of the solenoid is switched off.

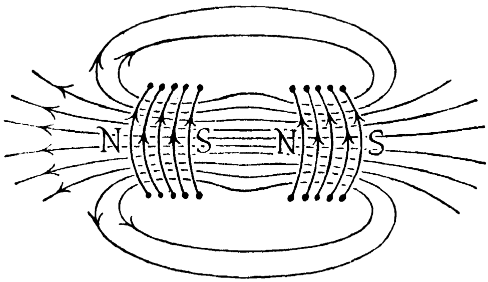


FIG. 108.

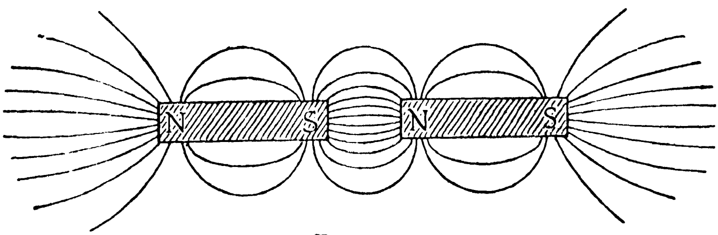


FIG. 109.

Fig. 108 shows neighbouring solenoids carrying like currents, and Fig. 109 shows neighbouring bar-magnets with unlike poles adjacent.

Fig. 110 shows neighbouring solenoids carrying unlike

currents, and Fig. 111 shows neighbouring bar-magnets with like poles adjacent.

Figs. 108 and 110 show clearly why like currents in solenoids give adjacent unlike "poles," and unlike currents give adjacent like "poles." They also show why it is as reasonable to speak of the poles of a solenoid as to speak of the poles of a magnet.

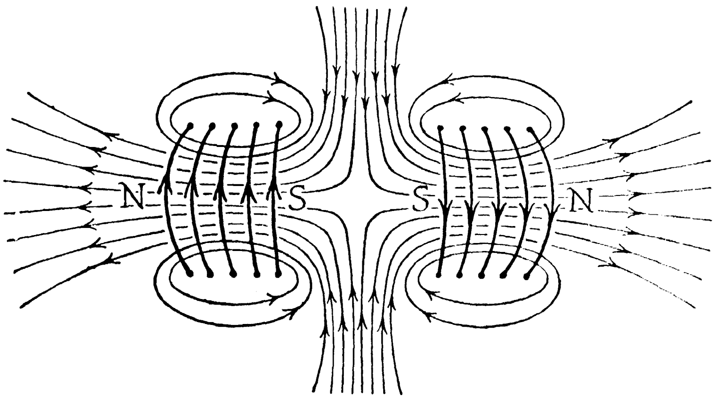


FIG. 110.

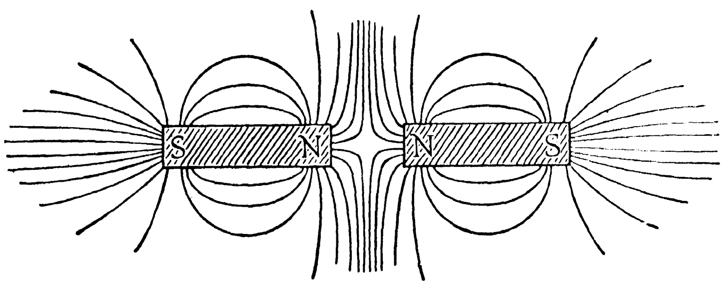


FIG. 111.

Fig. 112 shows a piece of soft iron between unlike poles of two solenoids, and Fig. 113 shows two pieces of soft iron between unlike poles of two bar-magnets.

Fig. 114 shows the lines of force inside an endless ring solenoid. These resemble the lines of force inside a magnetised iron anchor-ring; but such lines cannot be plotted directly because they are all inside the iron.

These figures show clearly why it is reasonable to suppose that the magnetic elements in a piece of magnetic material

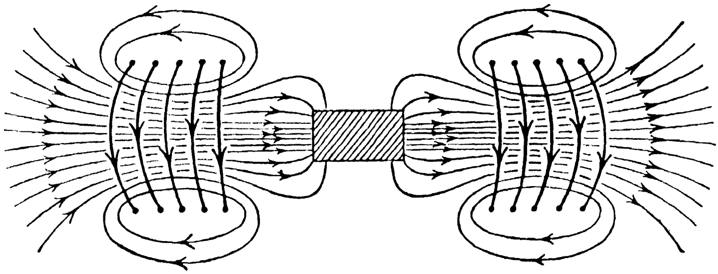


FIG. 112.

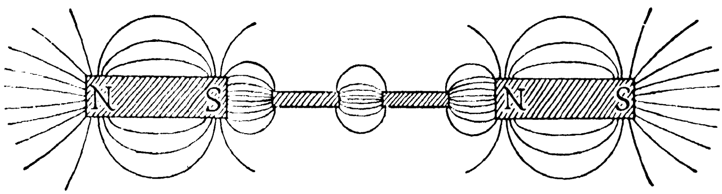


FIG. 113.

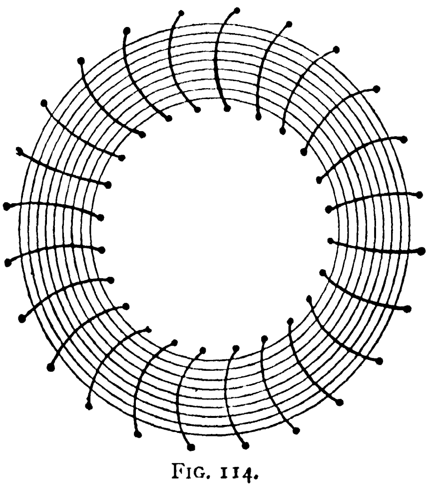


FIG. 114.

are small circular currents, or revolving electrical charges. If the orbits of such revolving charges are turned, when the

material is magnetized, so that their axes lie along the direction of magnetization, the lines of force through these circuit-elements must take the form actually observed.

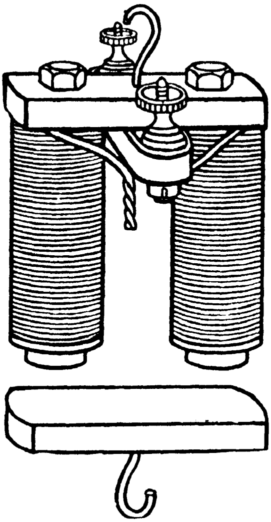


FIG. 115.

The experiments show clearly that the hypothesis of a magnetic material, or of the existence of a magnetic pole as a real entity rather than a mathematical convenience, is quite unnecessary. The hypothesis of the magnetic field as one of the manifestations of a current, or of a revolving charge, is enough to account for the facts in a general way.

Electromagnets

If a soft iron bar is put inside a solenoid carrying a current, it is of course strongly magnetized at once in the direction of the field along the

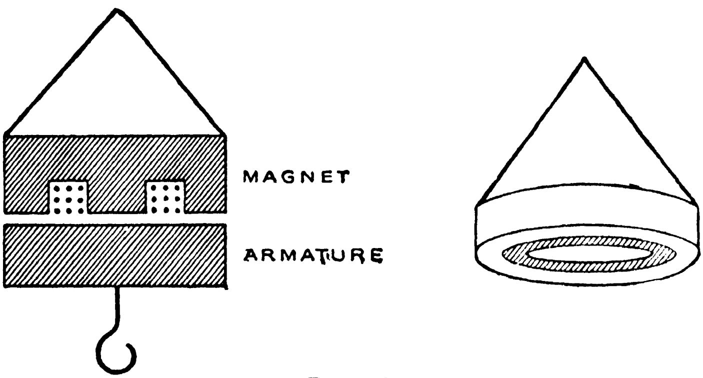


FIG. 116.

axis. The number of lines of force through the solenoid is thus enormously increased.

Actually when these lines penetrate the iron they are called

lines of induction. But they still continue as lines of force at the point where they leave the iron.

The ordinary electromagnet has the horse-shoe form shown in Fig. 115, but the most useful kind for lifting large weights is shown in Fig. 116. The coils are wound in a groove in a soft iron cylinder.

The effective discussion of electromagnets, and the behaviour of soft iron in magnetic fields, needs some mathematical foundation, and appears in Chapter V of Part II.

Force on Soft Iron in a Magnetic Field

All pieces of soft iron are attracted by all magnets. How is this to be explained in terms of the relations between magnetic poles and magnetic fields?

A magnetized piece of soft iron has equal and opposite

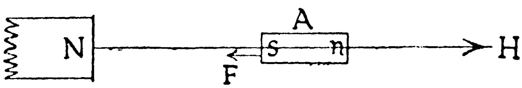


FIG. 117.

poles. Why do they not experience equal and opposite forces, so that the resultant force on the iron is zero?

This simple problem is sometimes misunderstood, and an explanation may be worth while.

A piece of soft iron experiences a force only in a non-uniform magnetic field. In a uniform field (such as the earth's field over a small area) it experiences no translational force, though there is a small couple.

In Fig. 117 let A be a small piece of iron which is near the pole N of a large permanent magnet. Let H be the field due to the pole N. H clearly diminishes as we get further from N.

A is magnetized by H with equal and opposite poles, *s* near N, and *n* away from N.

The attraction of the nearer pole *s* is greater than the repulsion of the further pole *n*, because H has diminished in getting further from N.

Thus the net force F acts toward N, whatever its polarity.

Pieces of soft iron in a magnetic field exhibit other properties which can be demonstrated by mapping the lines of force.

The lines of the field appear to be sucked into the iron as in Fig. 118, as if it were easier for them to pass through the iron

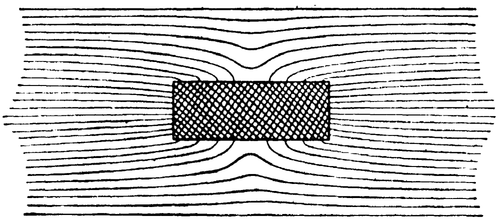


FIG. 118.

than through the air; as if the iron were more *permeable* than the air. This property does, in fact, follow from the permeability, which is considered in detail in Chapter V, Part I. An iron shell also serves as a magnetic shield.

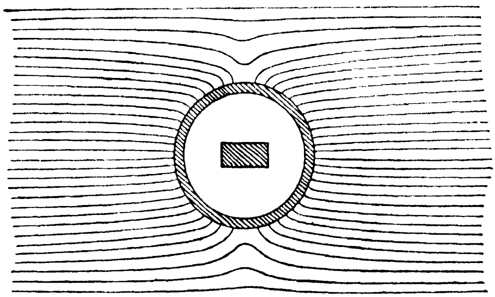


FIG. 119.

A magnetic field does not in general penetrate an iron shell appreciably. This property of iron can be demonstrated by plotting the lines of force in the neighbourhood of an iron ring in a magnetic field. A piece of soft iron inside the ring is not appreciably magnetized. The arrangement of the lines of force is as in Fig. 119.

Force on a Wire carrying a Current in a Magnetic Field

In Fig. 120, let NS be a long magnet whose poles may be regarded as concentrated at N and S. As S is far from N we may consider the effect of a local magnetic field on N alone.

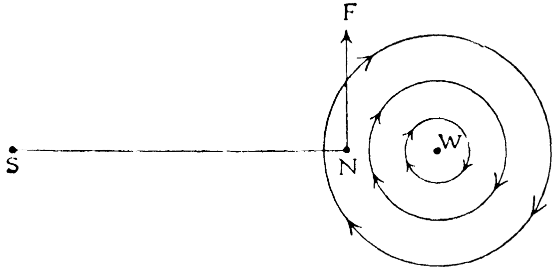


FIG. 120.

Let W be the cross-section of a wire through which a current is flowing vertically downwards into the paper. Consider that W is fixed, and N is free to move. The direction of the field due to W will be as shown by the little arrows on the circles. The direction of the force on N due to this field will be as shown by the arrow F.

Now let us suppose that the pole N is fixed, and the wire W is free to move. By Newton's 3rd Law, that Action and Reaction are equal and opposite, the force exerted by the pole on the wire is equal and opposite to that exerted by the wire on the pole. The situation, looked at from the point of view of N, is then as in Fig. 121. At the wire W there is a magnetic field H due to the pole N, in the direction of the arrow **H**, and a force in the direction of the arrow F.

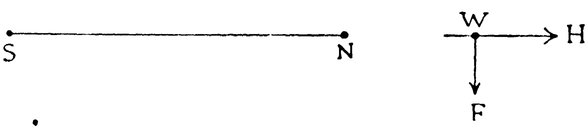


FIG. 121.

Now it is important to realize clearly that it is unprofitable to regard the pole N as causing the force F directly. The

pole N sets up the field H at the wire, and the presence of the field H causes the force F. The magnitude of this force can be simply deduced from the definition of the absolute unit of current.

By this definition we see that a current i flowing in one complete circle of wire of radius r cm. produces a field of $\frac{2\pi i}{r}$ gauss at the centre. This is deduced from the definition in Chapter VIII. Let a pole m be at the centre. It experiences a force $\frac{2\pi i m}{r}$ dynes.

If the pole be regarded as fixed, and the circle of wire as experiencing a force, this force is also $\frac{2\pi i m}{r}$. We should, however, regard this force as due to the field H due to the pole.

At the circumference $H = \frac{m}{r_2}$.

We may thus say that the force on a circle of wire of radius r is $2\pi i r$H dynes when the field is H perpendicular to the wire at every part of it.

It is reasonable to suppose that this force is evenly divided between all parts of the wire. The length l of wire is $2\pi r$; so we see that the force on a wire of length l cm., in a field H oersted, carrying a current i absolute units, is Hil dynes

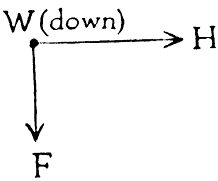


FIG. 122.

This proof is done properly in Part II, Ch. III, where calculus is available.

We have now evolved from first principles (using only the corkscrew rule) the relation shown in Fig. 122. If a current in W flows downwards through the paper, and a field H acts at right angles to the current along the paper, then the wire carrying the current will experience a force F, in the direction of the arrow F, at right angles to both H and W.

Fleming's Left-hand Rule

This relation has been expressed in many ways, of which the easiest seems to be Fleming's Left-hand Rule for a Con-

ductor carrying a current in a magnetic field. By this rule, if the First Finger of the left hand points in the direction of the Field (F—F—F), and the second finger points in the direction of the current, then the thumb (the strongest, and thus associated in the mind with force) points in the direction of the force acting. The three are held mutually at right angles with the first finger pointing straight out.

The writer has been enabled to remember this rule by a drawing he once saw but cannot now trace.

An L.C.C. tram-conductor stands in a large field in which are growing buttercups and magnets. His left arm is extended, and the first two fingers and the thumb are in the correct relative positions. At the tip of the second finger a current can just be seen. This drawing circulated in Field Survey Units in the Great War. The writer would be glad to rediscover it.

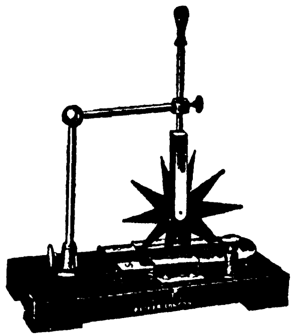


FIG. 123.

Barlow's Wheel

One of the most direct and striking ways of representing the action of Fleming's Left-hand Rule is the experiment known as Barlow's Wheel (Fig. 123).

A metal disc is fixed to revolve freely in a vertical plane, its lower edge in contact with a mercury bath between the poles of a strong electromagnet. There must be good electrical contact between the supports of the disc and the disc itself, through its bearings.

If a fairly large current is passed from the centre of the disc to the mercury bath, it will be compelled always to flow vertically downwards at right angles to the field between the poles of the magnet. Suppose the positive terminal of a battery is connected to the terminal on the left in Fig. 123. The current flows down each prong of the wheel in turn to the mercury bath between the poles, and out by the terminal on the right. By the left-hand rule the wheel will revolve in a clockwise direction (near pole of magnet is south).

Stretched Wire in a Magnetic Field

If a strong electromagnet is fixed in position with a retort-stand, and a piece of copper wire is fixed so that it is stretched moderately (but not very) tightly across the flat face of one of the poles of the electromagnet, the wire is found to jump sideways if a current through the wire is switched on while a current already flows through the electromagnet. It will also jump if the current through the wire already flows, and the electromagnet is switched on. This effect can be shown with

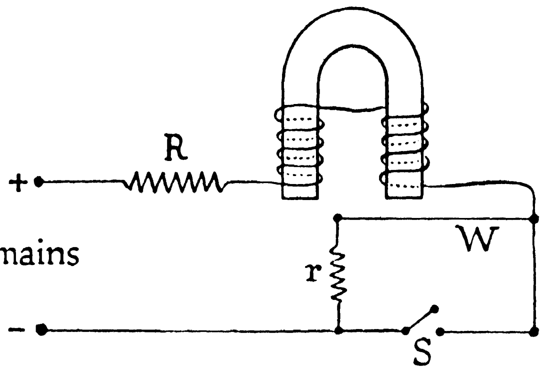


FIG. 124.

an ordinary large laboratory electromagnet carrying about 5 amps., and a current of about 5 amps. through the wire.

The experiment can be conveniently done with a circuit like that of Fig. 124.

R is a large resistance of the correct value to allow the desired current from the mains to flow through the electromagnet. The jumping wire W is of copper. r is a small coil of resistance-wire. S is a short-circuiting plug or switch. When S is shut it shunts W and r, and very little current flows through W. When S is opened, the current is diverted through W, and W jumps sideways.

This illustrates the principle of the Einthoven Galvanometer. See p. 403.

Couple on Coil carrying Current in a Magnetic Field

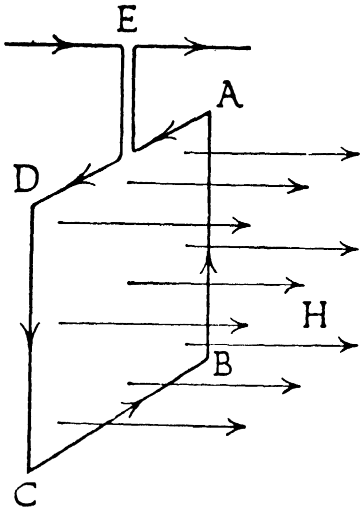


FIG. 125.

In Figs. 125 and 126 let ABCD be a rectangular coil carrying a current in a magnetic field.

In Fig. 125 let the field, shown by the arrows H, be perpendicular to the plane of the coil. In Fig. 126 let H be along the plane of the coil.

The current in the coil in each case is flowing in the direction DCBA as shown by the arrows.

In Fig. 125 the application of the left-hand rule to DC, CB, BA, and AD separately will show that the force on each arm acts outwards from the centre of the coil. The only possible effect of the field is thus to expand the coil. No couple acts. If either current or field were reversed the effect would be to contract the coil. If both were reversed it would be expanded. These statements can all be tested by Fleming's Left-hand Rule.

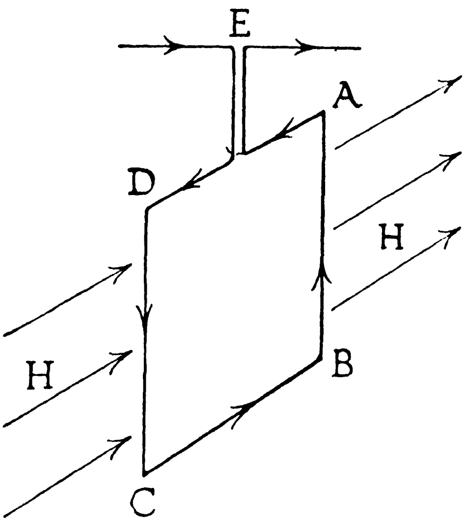


FIG. 126.

Now consider the effect on the coil in Fig. 126.

No force acts on AD or CB, since they are pointing directly along the field. DC experiences a force acting outwards from the paper, and BA a force acting into the paper. The coil thus experiences a couple which is anti-clockwise when observed from the top of the coil at E where the current-leads are. If either the direction of the field or the direction of the current through the coil is reversed, the couple will, of course, be reversed. If both are reversed, the direction of the couple is the same.

The general effect is that the couple always acts so as to turn the coil in such a direction that its plane is perpendicular to the field. In that position it stops in equilibrium. One may say that the coil always moves so that the maximum number of lines of force goes through it. It is possible to show (see Chapters II and III of Part II) that if a current of i absolute units flows through a coil of area A, parallel to a field H, having n complete turns, the moment of the couple is

$$\mathrm{H}i\mathrm{A}n \text{ absolute units of moment.}$$

[The absolute unit of moment is the moment of a force of 1 dyne acting at a distance of 1 cm. from the fulcrum.] If i is measured in amperes, the couple is

$$\frac{\mathrm{H}i\mathrm{A}n}{10} \text{ units.}$$

The couple is thus independent of the shape of the coil, but simply proportional to the area.

This couple acts in the moving-coil galvanometer.

The effect may be demonstrated with the two big coils. One is fixed, and the other freely suspended with its plane pointing at the centre of the fixed coil, along its axis. When the current is switched on, the movable coil turns round until its plane is parallel to the plane of the fixed coil; for in this position the plane of the coil is perpendicular to the field.

The effect may perhaps be better shown if quite a small coil, separately fed with current, is freely suspended in the field of one of the big coils.

Forces between Currents

If one of the big coils is laid flat on the bench, and the other laid on top of it so that the current in one goes clockwise and that in the other anti-clockwise, looked at from above, it will be found that the top coil leaps off as soon as the current through both is switched on. The effect is very striking, especially to people who do not expect it. It would have looked particularly like witchcraft a few hundred years ago.

If the currents in the coils are made to go in the same direction, both clockwise or both anti-clockwise, a strong attraction between the coils is observed.

These experiments demonstrate that parallel wires carrying currents in the same direction attract each other; and those carrying currents in opposite directions repel one another. They also demonstrate the close resemblance between a coil carrying a current and a bar magnet.

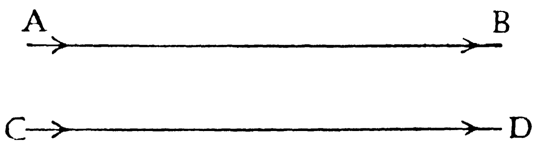


FIG. 127.

The fact that like currents attract and unlike currents repel can easily be deduced from Fleming's Left-hand Rule.

If AB and CD are two parallel wires carrying currents in the same direction (that of the arrows in Fig. 127) the field along CD due to AB will, by the corkscrew rule, be downwards into the paper.

By Fleming's Left-hand Rule the force on CD will be toward AB. So the wires attract each other.

Similarly, if the currents flow in opposite directions it can be shown that the wires repel each other. This can just be demonstrated for single wires if a large current (about 15 amps.) is available. If two wires (preferably part of the same circuit) are hung very close to one another (about ½ cm. or less apart) it will be found that they approach one another

if the currents are in the same direction, and separate a little if the currents are in opposite directions.

Induction of Currents

If a piece of wire (or anything that can carry a current) is moved across the lines of force of a magnetic field, an electromotive force appears in the wire. A current will therefore flow if the wire is part of a complete circuit.

This current can easily be shown to be increased if

(*a*) The wire is moved faster;
(*b*) The magnetic field is made stronger;
(*c*) The length of wire moving across the field is made greater.

These effects can easily be demonstrated by means of a wire solenoid of about two hundred turns, connected in series with a sensitive galvanometer. If a bar magnet is held in one hand and a solenoid in the other, and they are moved relatively to each other so as to make the solenoid surround the magnet, a current is seen to flow when, and only when, the two are moving relatively; and this current can be increased by moving the magnet faster, by replacing the magnet by a stronger one, or by replacing the coil by one with more turns.

These three effects are clearly the manifestation of a single law. In all of them the lines of force of the magnetic field are being cut across by the conductor. If the wire is moved faster, more lines are cut per second. If the field is made stronger, the number of lines available is increased, and more are cut per second. If the wire is made longer, more of it is occupied in cutting the lines, and more are cut per second.

The idea of the lines being "cut" by the conductor is, however, rather unsatisfactory, because we must have a complete circuit in order to have an induced current.

Though the *direction* of our induced current can most easily be found from the idea of cutting, the *magnitude* of it should come from the idea of "flux."

The "flux" through a circuit is the total number of lines of force threading the circuit. If the intensity of the field

were uniform all over the circuit, the flux would be the product of the area of the circuit and the magnetic intensity, since the magnetic intensity is the number of lines of force per square centimetre.

It can be shown experimentally that for any circuit the induced current is proportional at any instant to the rate at which the flux through the circuit is changing, in lines per second, and inversely proportional to the resistance of the circuit.

Clearly this means that the E.M.F. round the circuit is proportional to the rate of change of flux through it. It can be shown experimentally, and is deduced theoretically later in this chapter, that the induced E.M.F. in absolute electromagnetic units (1 volt = 10^8 electromagnetic units, as is shown in the chapter on Units and Dimensions in Part II) is actually equal to the rate at which the flux is changing in lines per second. Thus the flux through a circuit changes by N lines in t seconds, the induced E.M.F. is $\frac{N}{t}$ absolute units, or $\frac{10^{-8}N}{t}$ volts, if the flux is changing at a uniform rate.

If the flux is not changing at a uniform rate (and it seldom is) the problem can only be treated by calculus. This treatment appears in Part II, p. 347.

We may now formulate the first law of Electromagnetic Induction as follows :—

"*The electromotive force induced in a circuit through which the flux is changing is directly proportional (and, in absolute electromagnetic units, equal) to the rate of change of the flux in lines per second.*"

The direction of the E.M.F. is given by Fleming's Right-hand Rule, which may be regarded as the second law of Electromagnetic Induction.

If you hold your right hand so that the First Finger, representing Field, is pointed straight out, and the thumb, representing the direction of motion of the conductor across the field, is at right angles to the first finger, then the second finger held at right angles to each of the others gives the direction of the induced E.M.F. or current.

In the diagram (Fig. 128), if the dots represent cross-

sections of lines of force going *downwards* into the paper, and if the arrows AA give the direction in which the wire is being moved across the field, then the arrows BB give the direction of flow of the induced current.

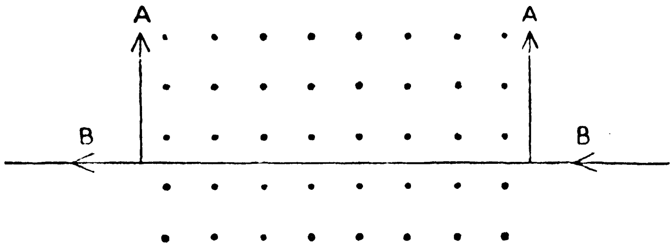


FIG. 128.

Magnetic Flux through a Circuit

If a complete circuit, such as a single coil of wire, is so arranged that there is a magnetic field going through it, then the total number of lines of force threading the coil is known as the *magnetic flux linked with the circuit.*

If Fig. 129 represented a coil of wire in the plane of the paper with a magnetic field perpendicular to it going downwards through the paper, the lines of force passing through where the dots show, then the magnetic flux linked with the circuit would be 4 units.

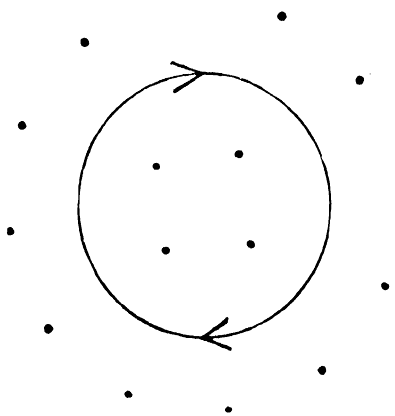


FIG. 129.

If there were two complete turns there would be 8 units, and, if there were n complete turns, $4n$ units.

If the magnetic flux became less, you could imagine this as represented by the movement of the lines of force across the boundary and out of the circuit.

Consider the left-hand bit of the circuit. The line of force

moves outward, from right to left. From the point of view, of Fleming's rule we may imagine that the line stays still, and that this bit of the circuit moves across it from left to right. By applying Fleming's rule you can easily see that the current is flowing in the direction of the arrow; and if you try the effect of moving a line of force out of any other part of the circuit, the movement will still make the current flow in the same sense.

Obviously, therefore, if the flux through the circuit diminishes, there will be a current induced in the direction of the arrows proportional to the rate at which the flux is changing multiplied by the number of turns in the coil; and, if the flux increases, the state of affairs will be the same, except that the current will flow in the opposite direction.

If now N_1 is the flux through the circuit at time t_1, and N_2 the flux at time t_2, the coil having n turns, the E.M.F. induced in the circuit (if the flux is changing at a uniform rate) is

$$n\left(\frac{N_2 - N_1}{t_2 - t_1}\right) \text{ absolute units.}$$

If the circuit has resistance R absolute units, the current i is given by

$$i = \frac{n}{R}\left(\frac{N_2 - N_1}{t_2 - t_1}\right).$$

The unit of flux, a single line of force, is called the Maxwell. Obviously, since H, the intensity, is the number of lines per sq. cm., it follows that if a uniform field H acts across an area A, the flux N is given by

$$N = HA.$$

Let us consider a simple numerical example.

The intensity in an air-gap is 50,000 gauss.

A coil of area 40 sq. cm. having 2000 turns is moved out of it (or into it) so that the flux through the coil changes at uniform rate from zero to its maximum (or from maximum to zero) in two seconds. The resistance of the coil is 5 ohms. Find the current.

$$\begin{aligned} i &= \frac{2000}{5 \times 10^9} \times \frac{(40 \times 50{,}000)}{2} \\ &= 0{\cdot}4 \text{ absolute unit} \\ &= 4 \text{ amperes.} \end{aligned}$$

R is put in as 5×10^9 units because 1 ohm is 10^9 absolute electromagnetic units of resistance. This is explained in Chapter IX, Part II.

It may have been noticed that we can only deal with changes of flux at a uniform rate, and can therefore discuss no real problem. Real problems require calculus, and the mathematical discussion of them appears in Part II.

It may, however, be noticed that for a steady current the total charge Q circulating is given by

$$\begin{aligned} Q &= (t_2 - t_1)i \\ &= (t_2 - t_1) \cdot \frac{n}{R} \cdot \frac{N_2 - N_1}{t_2 - t_1} \\ &= \frac{n(N_2 - N_1)}{R} \\ &= \frac{\text{Total number of lines cut}}{\text{Resistance of circuit.}} \end{aligned}$$

This delightful result can be shown to be universally true. It is the fundamental equation of the ballistic galvanometer. The deflection of a ballistic galvanometer is proportional to the total charge passing. Thus it gives the number of lines cut, and hence the flux, directly if it is calibrated for charge and its resistance is known.

Suppose, for example, that a ballistic galvanometer gave a throw of 20 cm. for a charge of 1 microcoulomb, or 10^{-6} coulomb.

When a coil of 100 turns is placed suddenly in an air-gap so that all the lines crossing the gap are cut by the coil (or better if it is suddenly removed from the gap) a deflection of 15 cm. is obtained. The resistance of the galvanometer and coil circuit is 300 ohms. Find the flux.

$$\begin{aligned} \text{Charge passing} &= \frac{15}{20} \times 10^{-6} \text{ coulomb} \\ &= 0{\cdot}75 \times 10^{-7} \text{ absolute unit.} \\ \text{Resistance} &= 300 \text{ ohms} \\ &= 3 \times 10^{11} \text{ absolute units.} \\ \text{Now} \qquad Q &= \frac{nN}{R} \\ \text{So} \qquad N &= \frac{QR}{n} \\ &= \frac{0{\cdot}75 \times 10^{-7} \times 3 \times 10^{11}}{100} \\ &= 225 \text{ lines, or Maxwells.} \end{aligned}$$

Lenz's Law

Suppose in Fig. 130 the dots were the places where a magnetic field perpendicular to the paper sent its lines of force through downwards, and suppose you moved the wire in the direction of the arrows AA; it is easy to see that the induced current would flow in the direction of the arrows BB (right-hand rule), as before.

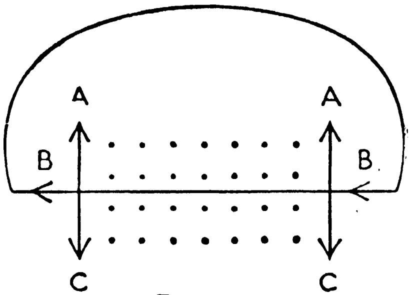


FIG. 130.

There will now be a force on the wire because it carries a current in a magnetic field, and by applying the left-hand rule you can see that the direction of the force will be that of the arrows CC, *opposing the motion which sets up the current.* It follows that, if a conductor moves across a magnetic field so as to have a current induced in it, the direction of the current will be such that the force on it opposes the motion which caused it.

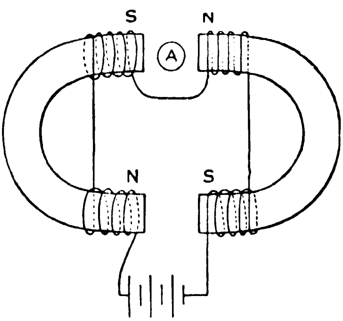


FIG. 131.

The last sentence is really a statement of Lenz's law, though rather a long one, and it was cast in such a form in order to show the law to be dependent upon the left-hand and right-hand rules, rather than a separate one.

The law, which may be regarded as the "Third Law of Electromagnetic Induction," may be stated:

"*Induced currents flow in such a direction as to cause a force opposing the motion that sets them up.*"

Lenz's Law can easily be demonstrated by arranging two electromagnets as in Fig. 131, and hanging, in the strong field between one pair of opposite poles, a copper cylinder A,

attached by a hook to a strong piece of cotton tied at the other end to some fixed support.

Turn the cylinder until the cotton has got enough twist to make the cylinder turn round quickly as soon as it is let go. When it is turning quickly, press the key so as to make the circuit of the electromagnet. The cylinder should very suddenly slow down almost to rest, and then continue to revolve slowly without stopping until the cotton is quite untwisted. If the circuit of the electromagnet is broken the cylinder should immediately speed up.

This phenomenon is sufficiently explained by Lenz's Law alone, since the induced currents set up in the cylinder must be in such a direction as to oppose its motion, but it is perhaps more satisfactory to work out the effect in detail.

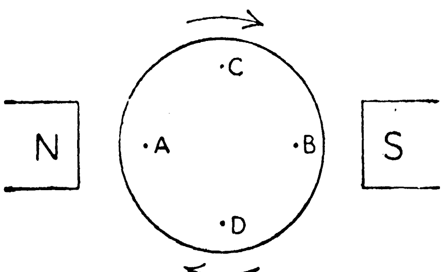


FIG. 132.

Suppose the direction of rotation of the cylinder, and the nature of the poles, are as shown in Fig. 132. Consider what is happening at the points A and B.

The right-hand rule tells us that a current is flowing downwards at A and upwards at B, but that at C and D no current at all is flowing, since there the motion is along the magnetic lines of force, instead of across them.

We have, then, only two currents to think about—the downward one at A and the upward one at B.

Consider now the force on A, according to the left-hand rule. First finger from N to S for the field; second finger downwards for the current; and the thumb points in a direction opposing the existing rotation. Try the effect on the upward current at B. Again the force opposes the present motion.

If one could carry a strong enough magnet on a car, Lenz's Law would provide a most beautiful method of braking, for it could produce any amount of resistance without locking the wheel.

The Five Laws relating Currents and Magnetic Fields

These laws, all of which have been considered in this chapter, may be briefly summarized as follows:

1. The Corkscrew Rule for the direction of the magnetic field round a straight current.
2. Fleming's Left-hand Rule for the direction of the force on a current in a magnetic field.
3. Faraday's Law of Electromagnetic Induction, that the induced E.M.F. acting round a circuit is proportional to the effective rate of change of flux through it. [By the "effective" rate is meant the actual rate multiplied by the number of turns threaded by the flux.]

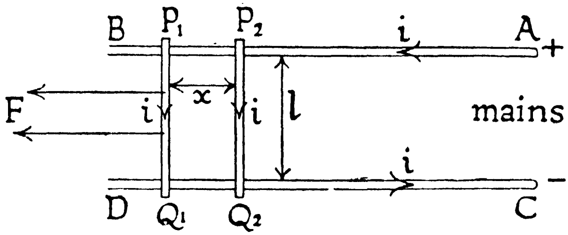


FIG. 133.

4. Fleming's Right-hand Rule for the direction of the induced E.M.F.
5. Lenz's Law that all induced currents flow in such a direction as to cause a force opposing the motion that set them up.

It is important to see that all the last four laws are dependent on the first, and can be deduced from it and the principle of the Conservation of Energy.

Law 2 is directly deduced from Law 1 at the beginning of this chapter.

Law 3 may be deduced as follows:

Let a wire of length l carry a current i at right angles to a magnetic field H.

In Fig. 133 let PQ, the wire, slide on rails AB, CD, from which the current is supplied. Let the field H act vertically upwards through the paper.

A force F ($= Hil$ dynes) acts on PQ, by Fleming's Left-hand Rule, in the direction of the arrows F.

Now let the wire be moved sideways a distance x against the force. Fx, or $Hilx$, ergs of work are done.

The change of flux N, the number of lines of force cut during the motion, is H $\times$ the area swept out.

Thus $N = Hlx$.

Thus $Hilx$ ergs of work are done in cutting Hlx lines of force.

We may thus say that if a conductor carrying a current i cuts a flux N, iN ergs of work are done.

The energy of the system has not changed. Where, then, has the work gone?

The only possible place for it is in the heat produced by i in flowing along l. Since i is unchanged an additional E.M.F. must have been acting between the ends of l, which produced an extra iN ergs of work in the time in which the motion happened. Let this time be t.

Then $\frac{iN}{t}$ extra ergs of work were produced per second, if the work was done at constant rate.

Thus if the extra E.M.F. was E

$$Ei = \frac{iN}{t}$$

$$\text{hence} \qquad E = \frac{N}{t}.$$

Thus an E.M.F. must be caused in a conductor moving across a magnetic field equal to the rate at which lines of force are being cut; or the E.M.F. in a circuit must be equal (in absolute units) to the effective rate of change of flux through the circuit.

It will be seen that the direction of the induced E.M.F. agrees with Law 4, and that a combination of Laws 2 and 4 must inevitably produce Law 5 in all cases.

Hence a great enough genius could have deduced all the other laws from Law 1 by theoretical methods alone.

Induction between Currents

We have seen that, if a wire and a magnetic field are moving with respect to one another, a current is set up in the wire.

But we do not need a magnet to produce a magnetic field, since every current has a magnetic field round it. When the current is flowing steadily, the magnetic field is stationary with respect to it, but when the current starts or stops, or changes for any other reason, its magnetic field changes, and may be regarded as moving.

The possibility of inducing a current without a magnet may be easily demonstrated with two wire solenoids of about two hundred turns each, one of which can fit into the other. If a steady current is sent through one, and the other is connected in series with a sensitive galvanometer, a deflection should be observed when the smaller one is being slid into the larger. It is rather more convenient to make the smaller one carry the current, to simplify the carrying out of the next experiment.

If some pieces of unmagnetized soft iron are now slid inside the smaller coil when the larger coil is already in position, a big deflection on the galvanometer should be observed, because the soft iron is causing the appearance of an enormous number of new lines of force, all of which may be regarded as cutting the wires of the solenoid in the process of arriving.

When the soft iron is again stationary, whether outside or inside the coil, the induced current through the galvanometer will again cea e to flow, of course; but, if the outer solenoid is again moveu when the inner one contains its soft iron core, the galvanometer deflection for a given rate of movement will be enormously greater than before the arrival of the iron core.

Starting and Stopping Currents

Two circuits, near enough to one another for the change in magnetic field, due to the current in one to induce a measurable current in the other, are known as "coupled circuits" (Fig. 134).

[The wireless enthusiast will recognize that his grid and anode circuits, in an ordinary magnetic reaction single-valve detector, are coupled circuits.]

Let us think about what happens in two coupled circuits in their simplest form when the steady current through one of them is started or stopped.

AB is a straight wire which is part of a circuit containing a cell V and a tapping-key K.

CD is a straight wire in series with a sensitive galvanometer G. AB and CD run parallel near one another.

When the current is flowing steadily from A to B, the magnetic lines of force due to it in the neighbourhood of CD are going downwards into the paper.

[The right-hand rule for steady currents and magnetic fields gives this at once.]

Before the current started flowing, there was no magnetic field. We may thus consider that, when the current started up, a magnetic field (having its intensity downwards through

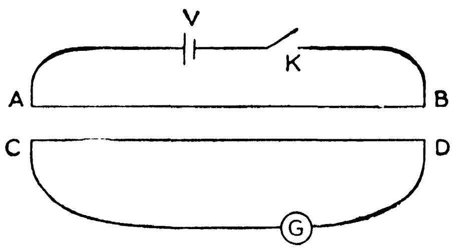


FIG. 134.

the paper) rushed out sideways from AB past CD. Since our rule for induced currents regards the magnetic field as stationary and the conductor as moving, we can suppose that CD was moving towards AB across a field downwards through the paper.

Try the right-hand rule for induced currents. First finger points downwards through the paper for the field; thumb from CD to AB for the motion across the field; and the second finger for the induced current points from D to C.

So a current starting up from A to B produces an induced current flowing from D to C.

Similarly, a current from A to B, when stopping, will cause the lines of force to close in to AB again, and so set up an induced current from C to D.

Up to now the situation is very easy to understand; but that is because we have entirely left out two other effects we ought to have considered.

We have seen that the change of current in AB causes a current in CD; but the starting of the current in CD equally causes a current in AB in addition to the original one, or rather makes the original current start at a rate different from that which it would have had if CD had not been there.

But now the induced current in CD is not quite what we originally thought it was, because the inducing current in AB has itself been affected by CD.

Thus the induced current in CD is affected, and produces a second change in the AB current, which produces a second change in the CD current, and so on. The situation is out of our mathematical control without calculus, and is dealt with in Part II.

Now let us imagine that AB and CD are brought nearer and nearer together till they coalesce. The CD current is not destroyed. It still opposes AB at starting, and acts in the same direction when AB is diminishing.

Thus there exists a tendency in every circuit for an opposing E.M.F. to act when the main current is growing, and an assisting E.M.F. to act when the main current is diminishing. Every current in a conductor thus behaves as if it had inertia.

Self-inductance and Mutual Inductance

The " inertia " of a current in a circuit, because it opposes its own changes, is called the Self-Inductance of the circuit.

It is obvious from Faraday's Law that the self-inductance is proportional to the algebraical sum of the number of times the lines of force from the circuit cut the circuit when they appear. Thus a solenoid has considerable self-inductance (which is enormously increased by putting some soft iron in it) because the lines of force from each turn can cut all the others as they move outwards, and the lines down the centre thread all the turns. If a piece of soft iron is put inside the solenoid, the number of lines of force threading the circuit for a given current is enormously increased.

If a circuit is wound as in Fig. 135, its self-inductance is very small, because the E.M.F.s induced in the two halves of each kink oppose one another, or (from another point of view) the lines of force due to the two halves of each kink circulate in opposite directions.

The large self-inductance of a solenoid is due to the mutual support of all the fields and induced E.M.F.s.

The unit of self-inductance may be regarded in two ways, which are easily seen to be equivalent.

The absolute unit of self-inductance is the self-inductance of a circuit in which 1 *line of magnetic force is linked with the circuit when* 1 *unit of current flows in it. It is also the self-inductance of a circuit in which a rate of change of current of* 1 *unit per second induces an E.M.F. of* 1 *unit.*

[Since we have seen that the induced E.M.F. *is* the rate of change of flux effectively linked with the circuit, these definitions are obviously equivalent.]

The corresponding definition for practical units is best expressed in the second form.

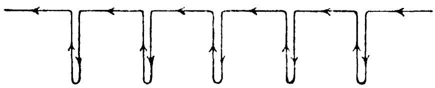


FIG. 135.

"*A circuit has a self-inductance of* 1 *henry when a rate of change of current of* 1 *amp. per second through it causes an induced E.M.F. of* 1 *volt.*"

Obviously if a rate of change of $\frac{1}{10}$ unit of current produces an E.M.F. of 10^8 units of E.M.F., the self-inductance must be 10^9 absolute units.

So $\quad 1 \text{ henry} = 10^9$ absolute electromagnetic units.

The units of mutual inductance, the inductance between coupled circuits, are of the same nature.

The absolute unit of mutual inductance is the mutual inductance of two circuits such that the passage of unit current in one causes the linkage of one line of force with the other. It is also the mutual inductance of two circuits in which a rate of change of 1 *unit per second, in one, produces an E.M.F. of* 1 *unit in the other.*

The mutual inductance of two circuits is 1 *henry when a rate of change of current of* 1 *amp. per second in one circuit causes an E.M.F. of* 1 *volt in the other circuit.*

Demonstrations of Self and Mutual Induction

The most direct demonstration is obtained by comparing the behaviour of two exactly similar solenoids, carrying identical currents when their circuits are broken.

If we connect our solenoids as in Fig. 136, either of them may be put in series with the rest of the circuit. As the inductance of a circuit without iron is very small compared to that of a circuit with iron, we may suppose that one circuit has resistance R only, and the other resistance R + inductance L.

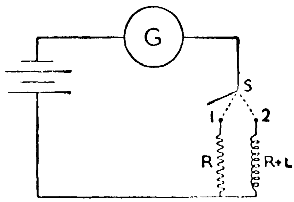


FIG. 136.

The only difference between the two circuits is that the solenoid in the inductive one contains soft iron, and the other solenoid does not. It will be found, however, that no spark, or only a small one, is seen when the non-inductive circuit is broken, but there is a large spark when the inductive circuit is broken. This spark is, of course, due to the leaping of the gap (a breakdown of its ionization) by the induced E.M.F.

Electromotive forces as big as 1000 volts can easily be produced when the circuit of a small electromagnet is broken, and quite an ordinary induction coil will give electromotive forces of 100,000 volts.

The inductive effect of a change of current may also be shown with this apparatus. When the switch *s* is in position 1 the galvanometer is in series with a resistance of small inductance. If the galvanometer is more or less dead-beat it goes straight to its final position when the switch is put to 1. When the current is switched off the galvanometer goes straight back to zero.

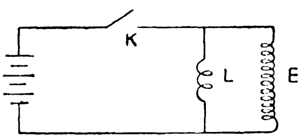


FIG. 137.

When the switch is in position 2, the galvanometer is in series with an equal resistance having large inductance. When the switch is put on the galvanometer goes more slowly to its final value; and when the current is switched off the galvanometer takes longer to get back to zero.

Self-induction may also be demonstrated by the apparatus

of Fig. 137. L is a lamp, E an electromagnet. Current is driven through both lamp and electromagnet when the switch K is closed. When K is opened, the breakdown of current in E causes an electromotive force round the circuit EL which produces a momentary bright flash in L.

Mutual induction is easier to illustrate. Perhaps the best illustration is the original experiment by which Faraday discovered Electromagnetic Induction. Two coils are wound round an iron anchor-ring, as in Fig. 138. A galvanometer G is in series with one coil, a battery B and a tapping-key K are in series with the other. When K is opened or closed, the anchor-ring is magnetized or demagnetized. Any flux in

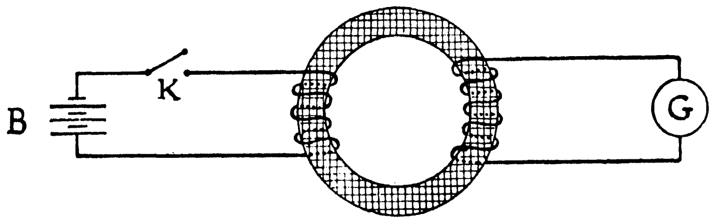


FIG. 138.

the ring is linked with the second coil. Thus any change of current in the KB circuit causes a change of the flux linked with the G circuit, and shows a current through the galvanometer.

No safety devices are included in the simple diagram shown. Of course the experimenter must see, by having suitable resistances in the circuits, that too great a current is neither taken from the battery nor sent through the galvanometer.

The coupled coils, in the ordinary swinging-coil method of applying reaction in wireless sets, give an excellent example of mutual induction. There is a certain amount of linkage of flux between the coils, and the presence of an oscillating current in one induces another oscillating current of the same period in the other.

Self-induction is illustrated in wireless apparatus by the high-frequency and low-frequency chokes. They produce back E.M.F.s, for a changing current, which can be arranged to act just long enough to stop an oscillatory current completely. The theory of chokes is discussed in Ch. VIII, Part II.

CHAPTER XIV

SIMPLE APPLICATIONS OF ELECTROMAGNETISM AND ELECTROMAGNETIC INDUCTION

Electromagnets—Buzzers—Telegraphs and Relays—Moving-coil Loud-speakers—Transformers—The Telephone—The Induction Coil—The Magneto-set for Cars—Other Applications.

Electromagnets

THE most interesting kind of electromagnet is that used in a laboratory for producing very strong fields over a small area. Such an electromagnet is shown in Fig. 139. The area of the pole-pieces is small compared with the area of cross-section of the core, in order to get the maximum possible flux-density in the gap. This instrument can produce fields as large as 50,000 gauss. Electromagnets for lifting large masses of scrap-iron out of rubbish generally take the form of Fig. 116. Electromagnets used in hospitals for picking magnetic dust from the eye have single poles, well separated, of the general shape of one of the poles of Fig. 139.

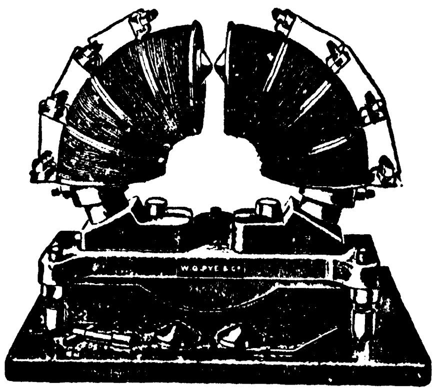


FIG. 139.

Buzzers

The ordinary electric bell or buzzer is operated by a small electromagnet as shown in Fig. 140.

A battery is connected through a tapping-key K to two terminals A and B. The circuit from A goes along a metal spring AT, having a striker T at the far end near the bell C. When no current is passing, AT rests in contact with a screw S, from which runs a wire round the field coils of the electromagnet to the terminal B. A soft iron armature D is fixed to the spring AT, just opposite the electromagnet.

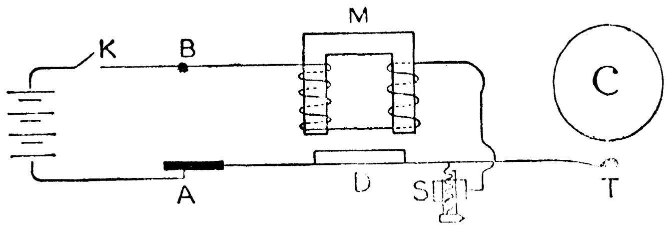


FIG. 140.

As soon as the key K is pressed, a current flows round the coils of the electromagnet, and it attracts the armature D. The whole spring AT moves, so that T hits C (giving a ring), and the circuit is broken at S.

The moment the circuit is broken M ceases to be a magnet, and the spring jumps back and remakes the magnet, so that the whole process starts all over again, and so continues till the switch K is opened.

In order that the spring may do work, it must receive energy. The inductive spark at break makes this possible, by allowing the electromagnet to do work for a short time after the spring has begun its return journey. If it were not for the self-induction in the circuit the spring would not be able to do external work.

Telegraphs and Relays

Telegraphs and Relays are worked as shown in Fig. 141. A " Relay " is a piece of apparatus which allows a large force to be controlled from a distance by a small one. The principle of an electric relay is that a small electromagnet, actuated from a distance by a small current, completes a circuit through which a large current can flow.

In Fig. 141 the electromagnet A makes the contact D, allowing the battery F to actuate the second, and much larger, electromagnet E. This part of the apparatus thus forms a relay.

The telegraph apparatus works as follows:

An electromagnet A is actuated by the current in the incoming line. It then attracts an armature B, which is held back, when not attracted, by a spring C. The armature B thus causes a contact D to close the circuit of a second, much stronger, electromagnet E, actuated by current from the battery F. E then attracts an armature G, held back, when not attracted, by a spring H, so that a pen K makes contact with some paper L, which is driven over rollers by clockwork.

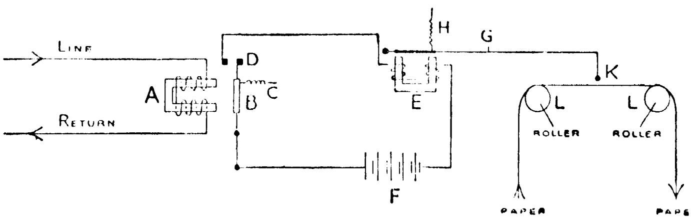


FIG. 141.

The system of dots and dashes in the original line-and-return system is thus represented by inked dots and dashes on the paper.

Moving-coil Loud-speakers

A recently invented and very interesting use of the electromagnet is in the moving-coil loud-speaker. The electromagnet used is of the form known as the pot-magnet, having one pole in the centre on a cylindrical soft iron core, and a circular air-gap separating this pole from the other, which forms a complete ring concentric with the cylinder on which the first pole is located. If the air-gap is narrow there is a strong field in it whose lines of force radiate from the centre of the core, and go straight across the gap. A coil of wire, to which a light stiff paper or composition cone is

rigidly fixed, is attached so that it is free to move in and out in the gap without touching either side (see Fig. 142).

The cone is held partly by the surround, and partly by other methods. One of the chief difficulties is to get the gap as small as possible, while keeping the coil accurately centred throughout the whole of its motion.

The interaction between the magnetic field in the air-gap and the oscillatory speech-currents in the coil cause the coil to move in and out with frequencies equal to the frequencies of the speech-currents.

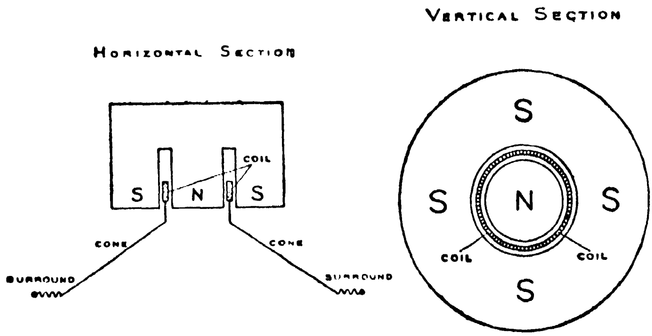


FIG. 142.

Transformers

A transformer consists simply of two coils of wire wound round an iron core.

If an alternating current flows through one of the coils (called the primary), a strong alternating magnetic flux is set up in the iron core.

This flux is linked with both coils, so that an alternating E.M.F. of the same frequency is induced in the second coil (called the secondary).

Since, as we have seen, the induced E.M.F. is proportional to the rate at which the lines of force are being cut, it follows that the induced E.M.F. in the secondary is proportional to the number of turns of wire in it. It is, in fact, possible to prove that a delightfully simple relationship holds approximately in transformers for all practical purposes. If the secondary has n times as many turns as the primary, then the

alternating voltage in the secondary is n times the alternating voltage in the primary.

This relationship only holds, of course, if the iron of the core is soft enough (or the frequency of alternation slow enough) for the magnetization of the iron to "keep up with" the magnetizing field of the primary current.

This is why iron-core transformers can be used only on the low-frequency side of a wireless receiver, where the frequencies dealt with range from 50 to 10,000 (and in this range transformers have their faults, particularly at the extreme values). They cannot be used on the high-frequency side, where the alternations have a frequency of the order of 1,000,000.

The cores of transformers are not solid, but are laminated, *i.e.* built up of strips of soft iron insulated from one another, to prevent the circulation of induced currents in the core. Induced currents in the core would heat it up, do no good, and absorb a great quantity of energy.

Transformers are used for two main purposes; to obtain high-potential alternating E.M.F.s, and to change a low-voltage high-current supply to a high-voltage low-current supply, or vice versa.

The former type is called a potential transformer, and is usually a small instrument used for wireless, telephone, telegraph, or laboratory purposes. It is comparatively easy to design, since the current carried by the secondary is either negligible or non-existent. The secondary circuit is, in fact, often not closed at all.

The latter type is called a current transformer, and is usually large and quite exceptionally difficult to design or manufacture, since considerable currents have to flow in both primary and secondary coils.

The designer, among other difficulties, must minimize eddy-current and magnetic lag (hysteresis) losses in the core, and eddy-current and resistance losses in the coils themselves. He must see that coils and core have adequate opportunity for cooling, that neighbouring coils are never at a high potential difference, and that the insulation is everywhere up to its job.

The use of a current transformer is easy to explain. Heat-losses in a given resistance are proportional to the square of

the current. Thus, by halving the current, the heat-loss is divided by four.

Consider now the case of a town at the centre of a large area for whose electrical supply it is responsible, and take the case of an outlying village ten miles away, whose current-consumption averages 500 amps. at 200 volts. The voltage used in the village must not be above 200, or it would be dangerous, but the erection of a wiring system ten miles long, capable of carrying 500 amps., would be intolerably expensive.

If, however, at the supply town, the voltage is transformed up to 40,000 volts, the current (not allowing for losses) will drop to $2\frac{1}{2}$ amps., since

$$40,000 \times 2\tfrac{1}{2} = 200 \times 500,$$

and the erection of a wire to carry $2\frac{1}{2}$ amps. is a cheap and easy job. At the village the supply can be transformed down again to 200 volts.

The heat-losses in this case, for a given supply wire, would be reduced in the ratio $\left(\frac{2\frac{1}{2}}{500}\right)^2$, or $\frac{1}{40,000}$.

Thus the heat-losses can be practically prevented by having a good transformer arrangement.

There is a further advantage in the use of transformers. Every designer of an alternating-current supply system has three voltages to consider:

(1) The best voltage for generating.
(2) The best voltage for house-supply.
(3) The best voltage for long-distance supply of power.

These are usually entirely different, the best generating voltage being normally higher than the house-supply voltage, but below the long-distance supply voltage.

The long-distance supply voltage should, of course, for economy, be as high as the insulating properties of the weakest part of the circuit—usually the transformer—will allow it to be.

Thus, it is easy to see that the design of current-transformers is both a very difficult problem, and a problem of the highest importance in commerce.

The Telephone

The following is an outline of the principles of the simplest method of telephony.

The microphone, or transmitting, end of a telephone system must be some device for producing electric oscillations, of the same frequency as the oscillations in air-pressure which form sound-waves. The simplest device for the purpose depends upon the fact that carbon granules have an electrical conductivity proportional to their tightness of packing. If they are packed behind a diaphragm which is made to oscillate by the sound waves, then the resistance of the carbon granules will vary with the same frequency as the sound waves; and, if the granules are included in an electrical circuit, the current in this circuit will vary in step with the sound waves, and the

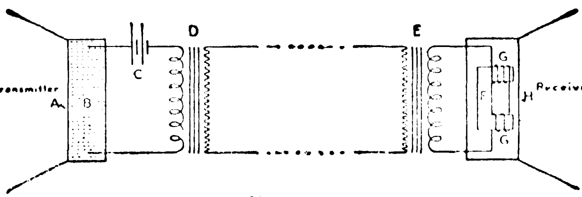


FIG. 143.

amplitude of the variations will (luckily) be approximately proportional to the amplitude of the sound waves. This oscillating current is sent through the primary of a step-up transformer. The secondary of this transformer is in series with the primary of a step-down transformer at the receiving end.

[This arrangement is exactly like the arrangement for carrying power across long distances in the form of alternating current.]

The secondary of the step-down transformer at the receiving end is in series with the windings of a small electromagnet whose poles are near a metal diaphragm. The attraction of this diaphragm by the electromagnet varies in step with the variations of the current, and the diaphragm, therefore, moves in and out, causing alternations of pressure in the air, and reproducing sound waves of the same wave-form as the

original ones which caused the current variations in the transmitter.

In Fig. 143:

A is the transmitting diaphragm.
B is the collection of carbon granules.
C is the battery which causes the primary current to circulate.
D is the step-up transformer at the transmitting end.
E is the step-down transformer at the receiving end.
F is the electromagnet at the receiver.
GG are the windings carrying the speech currents.
H is the receiving diaphragm.

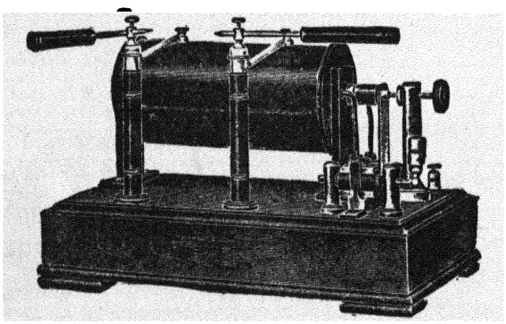


FIG. 144.

It must not be forgotten that this account of the telephone gives only the very barest outline of the simplest possible kind of telephony. The problems involved in the study of practical telephony are very difficult.

The Induction Coil

The Induction Coil (Fig. 144) is a piece of apparatus designed to produce very high voltages with negligible current. It is simply a combination of an electric bell with a transformer.

In the circuit shown in Fig. 145, the battery V sends a current through the primary coil via the platinum point P and the tapper T. As soon as the current passes, the field due to the primary coil makes the iron core a magnet, and

the tapper is attracted to it, so that T leaves P, and the primary circuit is broken at this point.

The magnetism of the iron core collapses as soon as the primary circuit is broken, and the spring S immediately drives T back against P, and makes the primary current again. This process continues indefinitely, as in the electric bell. Now the flux in the core is linked with the secondary, as well as the primary, so that, whenever there is a change in the primary current, at make or break, an induced E.M.F. appears in the secondary. If this E.M.F. is strong enough to break down the insulation of the air between A and B, the secondary terminals, a spark passes. If the spark is

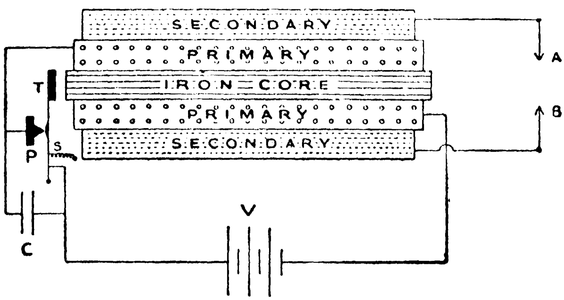


FIG. 145.

passing between sharp points, the distance it will jump gives a rough estimate of the E.M.F. in the secondary.

The E.M.F. is proportional to the gap it will jump, and is about 10,000 volts per cm.

Though the E.M.F. in the secondary appears at both make and break, the E.M.F. at make is quite negligible compared with that at break, and only the latter is considered.

It is important to remember that the induced E.M.F. depends on the rate of cutting of the lines of force, in lines per second. There are as many line-cuttings at make as at break, but, if there is not much sparking between T and P, the whole of the cutting is finished much more quickly at break.

The condenser C, shown in the diagram, is intended to minimize the sparking at break between P and S. It is worth while to minimize the sparking for two reasons. Firstly,

sparking means a continuation of the primary current, which in turn means a slow breakdown of it, which in turn means a small secondary E.M.F.

Secondly, sparking causes a burning of the points between P and S. These points are made of platinum in good induction coils, and tungsten in cheap ones, because these metals burn away slowest. The reason for sparking between P and S is that the primary coil, with its iron core, has a very considerable self-inductance, and produces a fairly high E.M.F. in the direction of the current at break. If there is a condenser across the break in the circuit, the energy of this induced

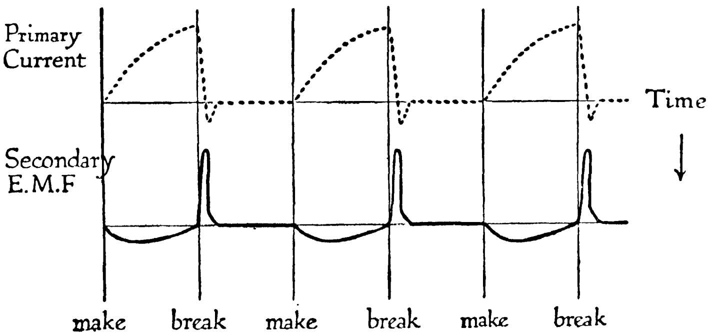


FIG. 146.

E.M.F. can exhaust itself in charging up the condenser, instead of driving the charge across the gap.

The curves of Fig. 146 show the kind of relation that usually exists between primary current and secondary E.M.F.

Notice that the *ordinate* of the secondary E.M.F. (full line) is at any point proportional to the *slope* of the primary current graph (dotted line) at the same moment.

The sudden sharp rise of secondary E.M.F. at break is the useful E.M.F. Since this rise is always in the same direction, the E.M.F. of an induction coil is for practical purposes unidirectional but intermittent.

The small reverse current in the primary, just after break, is due to the discharge of the current from the condenser. This current of course flows in the direction opposing the current from the battery. The sudden reversal of the

primary current at break helps the secondary E.M.F., so that the condenser is useful in two ways.

Ignition on Cars

A magneto uses a small dynamo as the primary of an induction-coil, and puts the secondary in series with each of the car's sparking-plugs in turn through a commutator.

The primary circuit includes the platinum points of a make-and-break system fixed on the end of the armature; and this circuit is broken by a cam during each of the periods when the secondary is in series with a sparking-plug. Every time the primary circuit breaks, the momentary E.M.F. induced in the secondary gives a spark at the plug then in the circuit.

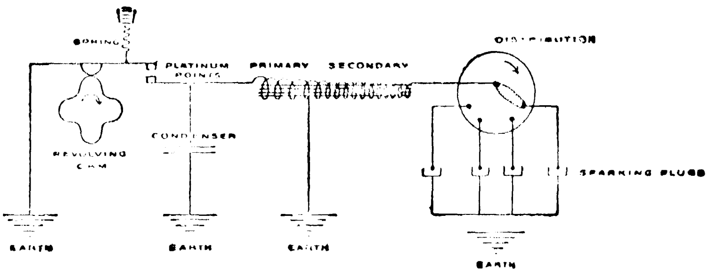


FIG. 147.

The time of this spark with respect to the motion of the piston can thus be accurately adjusted by shifting the cam-system of the make-and-break. Fig. 147 shows the electrical connections of the circuit, but gives no idea of the arrangement of the moving parts. One must take a magneto to pieces to understand it.

In coil ignition the dynamo part of the magneto is left out, and the secondary gets its E.M.F. as part of a battery-driven induction-coil.

The important things to watch in the design of a magneto are:

(1) That the spark is made to pass when the winding producing the primary current is in the position in which the lines of force are being cut by the secondary coil at *maximum* rate.

The spark of a retarded magneto is less powerful, because the induced current in the primary cannot be a maximum for more than one position of the armature. As the magneto is usually fully advanced, it is designed to give the biggest spark at full advance, and it therefore cannot give an equally big spark anywhere else.

(2) The platinum points should be clean, and the biggest separation neither too great nor too small. Dirty points and small separation cannot give a quick break and a big E.M.F. in the secondary. Too large a separation causes the break before the primary current has reached its full value, besides causing undue strain on the make-and-break.

(3) The carbon contacts of the distributor should be free from both insulating matter (like oil) which increases the secondary resistance, and conducting matter (like dirt) which allows partial short-circuits in the secondary.

(4) The gaps in the sparking-plugs should be neither too small, which would give a small spark, nor too large, which would put an undue strain on the internal insulation of the secondary.

Other Applications

Practically all electrical machinery is, of course, an application of electromagnetism or electromagnetic induction. The subject is so immense that no more than a brief outline is possible in a text-book of general electrical theory.

The following chapter gives a short account of Dynamos and Motors, and the chapter on Alternating Currents in Part II suggests the mathematical principles of the working of alternating-current machinery and instruments, which are less suitable for discussion in a general text-book than are direct-current instruments.

CHAPTER XV

DYNAMOS AND MOTORS

Principles of the Dynamo—The Commutator—The Armature—The Field-magnets—Series, Shunt, and Compound Windings—Lost Volts in Dynamo—Reverse Motor Effect—Numerical Problems about Dynamos—The D.C. Motor—Series and Shunt Motors—Increase of Speed for Reduced Field—Armature Reaction—Numerical Examples.

Principles of the Dynamo

If a coil of wire is turned in a magnetic field, an E.M.F. is induced in it equal (in absolute e.m. units) to the rate at which the flux through it is changing, in lines per second.

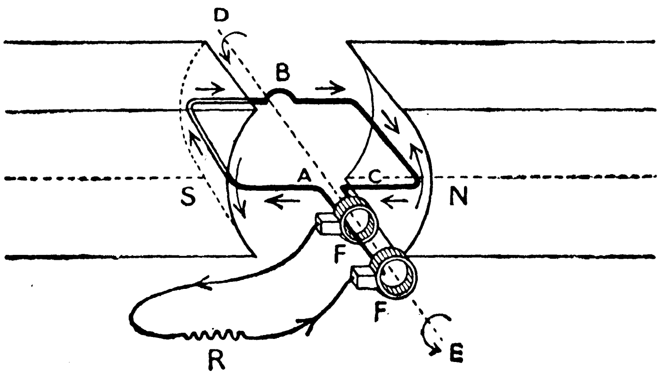


Fig. 148.

It is easy to see that this E.M.F. acts in one direction through one half of a complete turn, and in the opposite direction through the other half.

[Fleming's right-hand rule will show this at once.]

If by means of two *Slip-rings*, which make contact with the two ends of the coil all the time, we manage to complete the circuit of the coil so that the E.M.F. can send a current,

we have made a simple kind of alternating-current dynamo, or *Alternator.*

The diagram (Fig. 148) shows how to do this.

The coil ABC is turning in a counter-clockwise direction about the axis DE, between two poles N and S of a magnet. called the *Field-magnet.*

The ends of the coil are in sliding contact with the slip-rings FF, and, at the moment when the coil is in the position shown, a current is driven through the external resistance R in the direction shown by the arrows. The right-hand rule shows this. As the coil goes on turning, the total number of lines

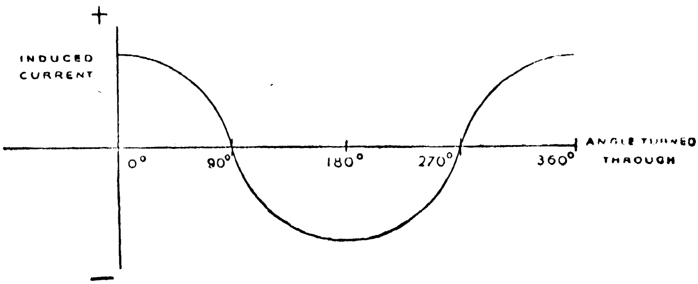


FIG. 149.

of force threading it increases, but the *rate* of increase becomes less, so the current becomes less.

As the coil passes through the up-and-down position, 90° away from the position shown, the number of lines of force threading it reaches a maximum, and then begins to decrease again.

Thus the *rate* of cutting the lines of force sinks to zero, and begins again in an opposite direction. The current flowing vanishes at the 90° point, and begins again in the opposite direction.

The current obviously reaches maximum value again, but in the opposite direction, when the coil is 180°, a complete half-turn, away from its present position. If the current is keeping up with the E.M.F. (which it does not necessarily do, as will be explained later), the graph of the current for a complete revolution will be as shown in Fig. 149.

The Commutator

By a very simple device, called a *Commutator*, the two halves of the current can be made to flow in the same direction.

In the diagram shown in Fig. 150 the two halves FF of the

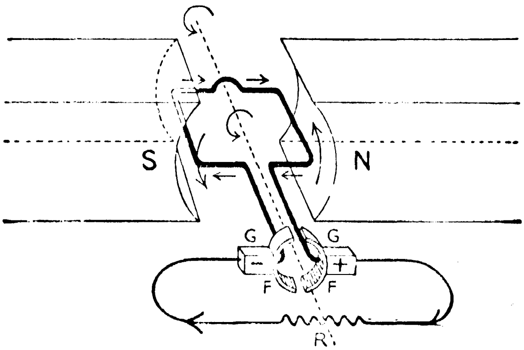


FIG. 150.

slip-ring are attached to the ends of the coil, and revolve with it (in a counter-clockwise direction, as before), but the *Commutator-brushes* GG are fixed and connected to the external circuit.

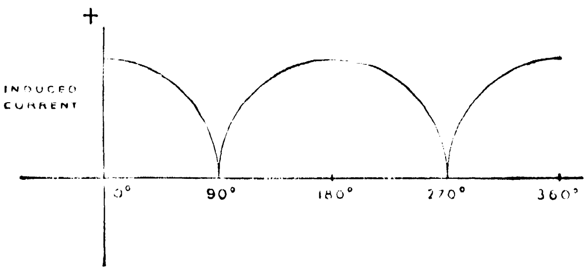


FIG. 151.

By applying the right-hand rule it is easy to see that whichever half of the slip-ring is in contact with the right-hand brush, this brush is positive and the other negative. *All* the current generated now flows through R in the direction of the arrow.

The effect of the commutator is thus simply to reverse the negative half of the current, and our current graph will now be as shown in Fig. 151.

This kind of current is in one direction, but is very jerky. Moreover the coil is only producing enough current to be

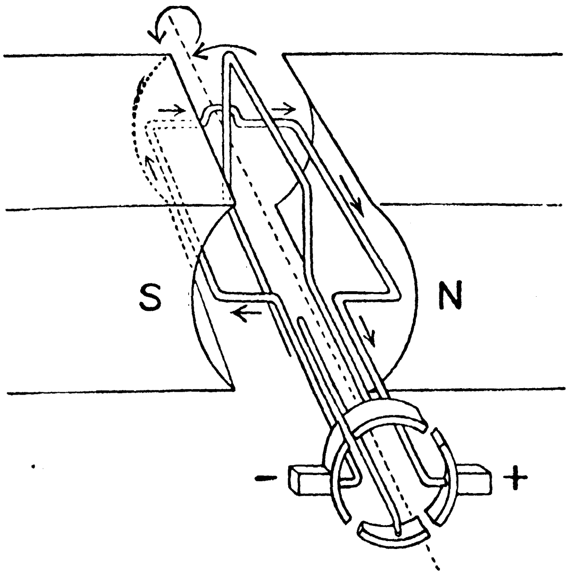


FIG. 152.

useful at the regions round 0° and 180°, and is producing very little near 90° and 270°.

The next improvement of our dynamo is to have two coils

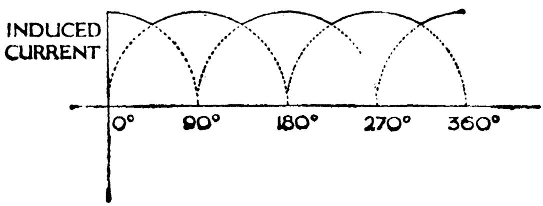


FIG. 153.

instead of one, and consequently four sections of the slip-ring instead of two. This arrangement is shown in the diagram (Fig. 152).

As before, the right-hand brush is always positive, and this time the current-graph will be as shown in Fig. 153.

This is much better. The current is never zero, or anything like it. It is true that the current has periodical variations in value, but we can make these variations smaller by increasing the number of revolving coils.

The kind of induced current obtained with sixteen equally spaced revolving coils is shown in Fig. 154.

This short account explains simply the fundamental principle of the alternator and the D.C. dynamo, and incidentally the principles of the D.C. motor (but not the A.C. motor) as well. If, instead of turning the coils round and taking out a current from the commutator-brushes, we send the current in through the commutator-brushes and see what happens, it is obvious, from Fleming's left-hand rule for the force on a conductor carrying a current in a magnetic field, that the system of coils will revolve, and that the direction of revo-

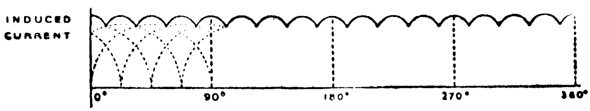


FIG. 154

lution will be *opposite* to that of the diagram if the direction of the current is the same.

Thus every dynamo is theoretically able to act as a motor, and, in fact, it will do so if given the chance.

The Armature

The *Armature* in a dynamo is a general name for the part where the E.M.F. is produced, and in the motor for the part carrying the driving current on which the driving force acts. In both, it consists of a number of coils of copper wire wound on a piece of soft iron which is free to revolve about its central axis. The reason for having a piece of iron, on which to wind the coils, is that it is important for the magnetic field between the poles of the main field-magnets to be as strong as possible; and the more iron there is between the poles, the stronger is the field. In our elementary diagrams, the iron of the armature was left out for clearness, and the armature in these cases consisted simply of the coils.

The coils wound on the armature are the *armature-windings* or *armature-coils,* and the current through them is the *armature-current.*

The Field-magnets

The magnets producing the magnetic field in which the armature revolves are called the *Field-magnets.* They are, of course, electromagnets (except in toys or shocking-machines). The coils wound on them are the *Field-coils,* and the current through them is the *Field-current.*

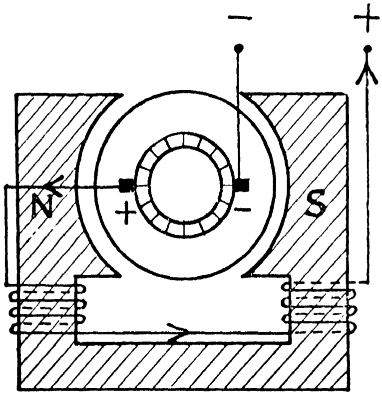


FIG. 155.

Series, Shunt, and Compound Windings

Dynamos are in general series-wound, shunt-wound, or compound-wound.

In a series-wound machine the field-coils are in series with the armature-coils, so that the same current flows through both, as in Fig. 155.

We may suppose that in most dynamos the core of the field-magnets is nowhere near saturation. So the strength of the field is roughly proportional to the field-current. Thus in all dynamos the average E.M.F. is proportional to (Field-current × Revs.), when " Revs." is short for the rate of rotation of the armature.

In the series dynamo the field-current and the armature-

current are identical. Thus the E.M.F. is proportional, for constant speed of revolution, to the current taken. In practically all cases of power-supply the first need is a constant E.M.F., independent of the amount of current taken. Consequently the series-dynamo is useless for almost all purposes. It is sometimes used for lighting arc lamps.

In the shunt dynamo, the field-coils are in parallel with the armature-coils, as in Fig. 156.

The E.M.F. is thus more or less proportional to the revs. and independent of the current taken. Actually it can be

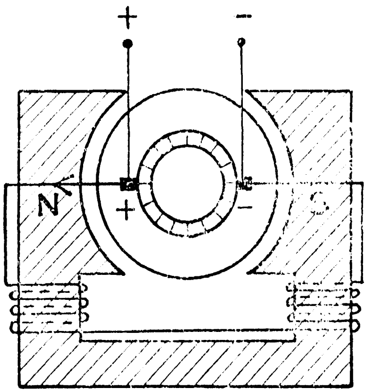


FIG. 156.

shown to diminish a little with increase of current. In general, thus, the shunt dynamo is useful.

The slight fall of E.M.F. with rise of current is, however, undesirable. It is counteracted in the "compound-wound" dynamo shown in Fig. 157.

We have noticed that in the shunt-wound dynamo the E.M.F. decreases a little with increase of current, and in the series-wound dynamo the E.M.F. increases a lot with increase of current. A judicious combination of the two can thus be made to keep the E.M.F. very nearly constant. Such a combination must obviously have many shunt-turns and a few series-turns in the field-coils.

In Fig. 157 a controlling resistance is shown in series with the shunt-windings. By this the field can be controlled. This *field-regulating resistance* is commonly used to vary the E.M.F.

Lost Volts in Dynamo

The supply-voltage (or P.D.) of a dynamo is bound to be less than the generated E.M.F. just as the supply-voltage of a storage-cell is less than its E.M.F. Both the dynamo and the cell have internal resistance, and the difference between

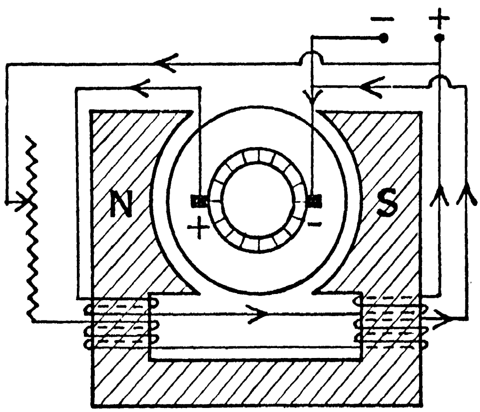


FIG. 157.

E.M.F. and P.D. is, for both, the product of current and internal resistance.

Thus a shunt-wound dynamo, with an armature-resistance of $\frac{1}{4}$ ohm and a field-resistance of 50 ohms, supplying an external current of 40 amperes at a P.D. of 200 volts, supplies 44 amperes altogether from its armature, allowing the extra 4 amperes for the field-coils. There is thus a P.D. inside the armature of $44 \times \frac{1}{4}$, or 11 volts. The E.M.F. must thus be 211 volts in order that the P.D. may be 200 volts.

The disappearing 11 volts are the "Lost Volts." It will be seen that they are directly proportional to the armature-resistance, which should therefore be as small as possible.

The power (11×44, or 484 watts in the case we are considering) is dissipated as heat in the armature-coils.

The conception of "Lost Volts" enables us to see why the E.M.F. and P.D. of a shunt-wound dynamo both fall with increase of current.

Suppose, in the case we have been considering, that the external current increased to 80 amperes, the revs. remaining the same. The lost volts would be increased so that the P.D. would fall; and the fall in P.D. would reduce the field-current, and thus reduce the number of lines of force cut per second. The E.M.F. would therefore fall and cause a further reduction in P.D.

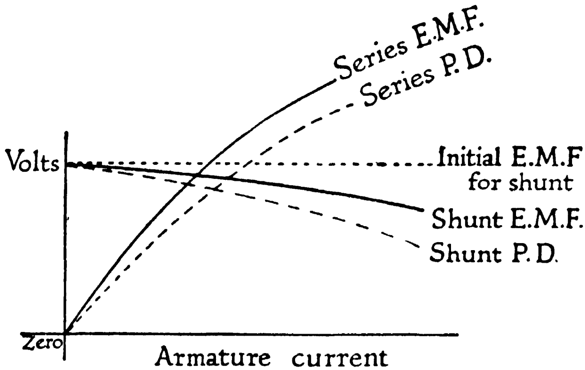


FIG. 158.

At the first moment of increase of current, the lost volts would jump from ($44 \times \frac{1}{4}$) to ($84 \times \frac{1}{4}$). The P.D. would thus fall to $211 - 21$ or 190. The P.D. would thus fall from 200 to 190, so that the field-current would drop from 4 amps. to 3·8 amps. The E.M.F. would thus drop from 211 to $\frac{19}{20} \times 211$, or 200·9 volts, and the P.D. would thus drop further, though only about half as far as before. P.D. and E.M.F. would thus drop alternately, each drop being about half the former one. Thus if the first drop of P.D. is v, the total drop is about $\left(v + \frac{v}{2} + \frac{v}{4} + \frac{v}{8} + \ldots\right)$, or $2v$. The final value of our P.D. would thus be about 180 volts.

In the case of a series-wound dynamo, the internal resistance is, of course, greater because it includes the field-coils. So the P.D. falls off more quickly with respect to the E.M.F.

The general shape of the curves relating E.M.F. and P.D. to armature-current for the series and shunt dynamos is shown in Fig. 158.

In all types (unless there are separately excited field-coils) the original build-up of E.M.F. when the machine starts is due to the residual magnetism in the field-magnets.

Reverse Motor Effect

By Lenz's Law a force will oppose the motion of the armature of a dynamo when it carries an induced current. Thus the work required to turn a dynamo increases with the current supplied. The Principle of the Conservation of Energy shows independently that this must happen. Moreover, since energy must always be lost in unproductive heating of both armature-coils and field-coils, as well as in overcoming friction, it is obvious that the power obtained from a dynamo must always be markedly less than the power supplied. The numerical value of the efficiency is usually taken as the percentage of power delivered to power supplied.

Numerical Problems about Dynamos

The general behaviour of dynamos is best understood if one works out a few simple numerical problems about them.

1. (*a*) A shunt dynamo supplies 33 amps. at 180 volts P.D. to an external circuit. Its armature-resistance is $\frac{1}{3}$ ohm, and field-resistance 60 ohms. Find the E.M.F.

$$\text{Lost Volts} = \left(\text{external current} + \frac{\text{P.D.}}{\text{Field-resistance}}\right) \times \begin{array}{c}\text{armature-}\\\text{resistance}\end{array}$$

$$= \left(33 + \frac{180}{60}\right) \times \frac{1}{3}$$

$$= 12 \text{ volts.}$$

$$\therefore \text{ E.M.F.} = (180 + 12) = 192 \text{ volts.}$$

(*b*) If the revs. were 3200, and they dropped to 3000 while the external current dropped to 27 amps., the field-current

being kept the same by altering the field-resistance, find the new P.D., and the amount of resistance which must be cut out of the field-regulator.

$$\frac{\text{New E.M.F.}}{\text{Old E.M.F.}} = \frac{\text{New Revs.}}{\text{Old Revs.}} \times \frac{\text{New Field-current}}{\text{Old Field-current}}$$

$$\frac{\text{New E.M.F.}}{192} = \frac{3000}{3200} \times \frac{3}{3}$$

$$\text{New E.M.F.} = 180 \text{ volts.}$$

Now $$\text{New P.D.} = \text{New E.M.F.} - \text{Lost Volts}$$

$$= 180 - (27 + 3) \times \frac{1}{3}$$

$$= 170 \text{ volts.}$$

$$\text{New Field-resistance} = \frac{\text{New P.D.}}{\text{Field-current}}$$

$$= \frac{170}{3}$$

$$= 56\tfrac{2}{3} \text{ ohms.}$$

So $3\frac{1}{3}$ ohms must be cut out of the field-resistance.

Problem 2.—A house-lighting set is connected up as in Fig. 159. Storage batteries of E.M.F. 100 volts are in parallel with a dynamo of constant E.M.F. 110 volts, and the combination is used to supply the lighting system. [This, a common arrangement, is known as a "floating battery" system.]

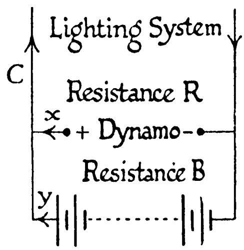


FIG. 159.

The armature-resistance of the dynamo is 0·08 ohm, and the internal resistance of the accumulator is 0·02 ohm.

The field-current of the dynamo may be neglected.

When the lamps are taking 100, 120, or 150 amps., find the supply P.D. and the current supplied by, or to, the battery.

The problem may be clarified if we solve it first for symbols. Let c be the external current, x the current supplied by the dynamo, and y the current supplied by the battery.

[If y is negative, clearly the battery is being charged by the dynamo.] R is the dynamo-resistance, B the battery-resistance. E is the E.M.F. of the dynamo, e the E.M.F. of the battery.

By Kirchoff's Laws

$$x + y = C.$$

Round the circuit containing dynamo and battery

$$Rx - By = (E - e).$$

The solution of these equations is easily seen to be

$$x = \frac{B}{R + B}C + \frac{E - e}{R + B}$$
$$y = \frac{R}{R + B}C - \frac{E - e}{R + B}.$$

The current supplied by the battery is thus positive or negative according as $(E - e)$ is less or greater than RC. In our case $E - e = RC$ when $C = \frac{110 - 100}{0{\cdot}08} = 125$ amps. Thus for 125 amps. external supply the dynamo will just supply all of it. For smaller currents the battery will be charged up. For larger currents the battery will help the dynamo.

The P.D. of the combination clearly $= E - Rx = e - By$. In our special case, when $C = 100$ amps.

$$x = \frac{0{\cdot}02}{0{\cdot}10} \times 100 + \frac{110 - 100}{0{\cdot}10}$$
$$= 120 \text{ amps.}$$
$$y = \frac{0{\cdot}08}{0{\cdot}10} \times 100 - \frac{110 - 100}{0{\cdot}10}$$
$$= -20 \text{ amps.}$$

Thus the battery receives 20 amps. charge.

P.D. of combination $= 110 - 0{\cdot}08 \times 120 = 100 - 0{\cdot}02\,(-20)$
$= 100{\cdot}4$ volts.

When $C = 120$ amps.

$$x = \frac{0{\cdot}02}{0{\cdot}10} \times 120 + 100 = 124 \text{ amps.}$$
$$y = \frac{0{\cdot}08}{0{\cdot}10} \times 120 - 100 = -4 \text{ amps.}$$

P.D. $= 110 - 0{\cdot}08 \times 124 = 100{\cdot}08$ volts.

When $C = 150$ amps.

$$x = \frac{0{\cdot}02}{0{\cdot}10} \times 150 + 100 = 130 \text{ amps.}$$
$$y = \frac{0{\cdot}08}{0{\cdot}10} \times 150 - 100 = 20 \text{ amps.}$$

E.M.F. $= 110 - 0{\cdot}08 \times 130 = 99{\cdot}6$ volts.

The effect of the floating battery should now be clear. As the external current rises from 100 amps. to 150 amps., the output by the dynamo only rises from 120 amps. to 130 amps., and the P.D. of the combination only sinks from 100·4 volts to 99·6 volts.

The dynamo is thus kept working at an almost constant current, and a still more constant E.M.F.

The D.C. Motor

The D.C. motor is, in principle, a dynamo working backwards, though actually the design is quite different. The direct current is made to pass in the correct direction through each separate winding of the armature-coils by means of the commutator. Each armature-coil is thus a coil carrying a current in a magnetic field, and it therefore experiences a couple which makes it revolve.

The motor acts also as a dynamo, and has a back E.M.F. induced in it by the motion of the armature-coils across the field. The difference between the applied P.D. and the back E.M.F. is effective in driving current through the armature-coils, and the work done by this current is dissipated as heat. Thus the Lost Volts for a motor are the volts which drive the armature-current!

Every D.C. motor needs a starting-resistance which must be gradually cut out as the motor speeds up. If there were no such device, much too large a current would go through the armature-coils at the beginning when the back E.M.F. had not yet appeared. As the back E.M.F. is proportional to (revs. $\times$ field), it is obvious that it cannot reach its full value till the motor has settled down at its permanent speed. Thus the starting-resistance cannot be all cut out until the final speed is reached.

The general relations of the motor should now be easy to see. We have obviously—

$$\text{Back E.M.F.} = \text{Applied P.D.} - \text{Lost Volts} \quad . \quad . \quad (1)$$
$$\text{Back E.M.F.} \propto \text{Field-current} \times \text{Revs.} \quad . \quad . \quad . \quad . \quad (2)$$
$$\text{Torque} \propto \text{Field-current} \times \text{Armature-current} \quad . \quad (3)$$

[*Torque* is the moment of the couple exerted by the motor. It is thus the vital property of the motor.]

By combining (2) and (3) we can get another relation—

$$\text{Torque} \propto \frac{\text{Back E.M.F.} \times \text{Armature-current}}{\text{Revs.}} \quad . \quad (4)$$

So that for constant revs. the force exerted by a motor is proportional to the back E.M.F. Thus if we are given the speed and armature-current of a motor we should try to get the back E.M.F. as big as possible—a somewhat surprising result. It may explain, however, how it is that we may regard the voltage driving the armature-current as "Lost Volts."

Let us consider a simple case to illustrate these relations.

Problem 3.—A shunt motor whose armature-resistance is $\frac{1}{3}$ ohm is run off 100-volt mains. When it is running light at 490 revs. per minute its armature-current is 6 amps. Find its back E.M.F. At what speed must it run if it is loaded till the armature-current is 24 amps.?

In what ratio has torque increased?

Since the field-current is 6 amps. at first, the effective volts or "Lost Volts" are $6 \times \frac{1}{3}$, or 2 volts. The back E.M.F. is thus $(100-2)$ or 98 volts.

If the armature-current is now increased to 24 amps., the Lost Volts increase to $\frac{24}{3}$ or 8 volts. So the back E.M.F. must drop to 92 volts.

Since back E.M.F. $\propto$ Revs. $\times$ Field, and the field has stayed constant because the applied P.D. is still 100 volts, the revs. must have decreased to $\frac{92}{98} \times 490$, or 460.

By relation (3) the torque has been multiplied by $\frac{24}{6}$, or 4. Thus in the new arrangement $\frac{1}{4}$ of the torque is being used to overcome friction.

Series and Shunt Motors

Since the field-current and armature-current are the same for the series motor, it is clear that torque $\propto$ (armature-current)^2. The faster the motor goes, the greater the back E.M.F., the smaller the armature-current, and the smaller the torque. It thus follows

(*a*) That a series motor can get an enormous starting-torque.

(*b*) That the general effect of a sudden increase of load is to slow down the motor very greatly, rather than to increase the current taken.

A series motor is thus useful for electric traction, but useless for driving machinery at constant speed.

In the shunt motor, since the field-current is constant, torque $\propto$ armature-current. A very slight slowing-down may cause an enormous increase in armature-current and torque. For example, if a motor running with a P.D. of 100 volts carries an armature-current of 10 amps. with an armature-resistance of 0·1 ohm, the back E.M.F. must be 99 volts. If the speed drops by $\frac{1}{99}$ of its value—only a shade over 1%—the armature-current, and consequently the torque, are doubled. Thus the shunt motor runs at almost constant speed.

The effect of a sudden increase of load is almost entirely to increase the armature-current, and only very slightly to slow down the motor.

It is thus specially useful for driving factory machinery where the load may vary greatly and suddenly but the speed must stay almost constant.

Shortly summarized, the difference in performance between the series motor and the shunt motor is that, if the load is doubled, the series motor takes $\sqrt{2}$ times as much current and runs at a little less than $\frac{1}{\sqrt{2}}$ of its former speed. The shunt motor takes nearly twice as much current, and continues to run at almost its former speed.

Increase of Speed for Reduced Field

The shunt motor has one very curious property. Reducing the field-current increases the speed, although, of course, it reduces the torque. This can best be illustrated by a numerical example.

Problem 4.—A shunt motor, with an applied P.D. of 100 volts, takes 10 amps., 8 amps. through the armature and 2 amps. through the field-coils. Its armature-resistance is 0·5 ohm. Its field-current is reduced to $1\frac{3}{4}$ amps. by increasing the

field-resistance, and at the same time the torque is reduced to $\frac{3}{4}$ of its former value. In what ratio is the speed increased, and what is the new armature-current?

When we first reduce the field-current, the immediate effect is that the back E.M.F. is decreased in the ratio $\frac{15}{16}$, so that it falls from 96 volts to 90 volts. The lost volts thus leap up from 4 volts to 10 volts, so the armature-current leaps at once from 8 amps. to 20 amps. This causes the torque to become nearly two and a half times as great for the moment, and the motor speeds up. The increase of speed increases the back E.M.F., which cuts down the armature-current and the torque again as the speed increases.

The speed at which the motor settles down again depends upon the load. If the new torque required is to be $\frac{3}{4}$ of the old one, then since torque varies as field-current $\times$ armature-current, and the field-current is reduced in the ratio $\frac{15}{16}$, it follows that the armature-current must be decreased in the ratio $\frac{16}{15} \times \frac{3}{4}$, or $\frac{4}{5}$. The new armature-current is thus $\frac{4}{5} \times$ 8 amps., or 6·4 amps., so the lost volts are 3·2 volts, and the back E.M.F. is 96·8 volts. So the new speed is $\frac{16}{15} \times \frac{96{\cdot}8}{96}$, or $\frac{242}{225}$ of the old speed.

The speed of a shunt motor is thus best controlled by a variable field-resistance. Another fact worth noticing about the shunt motor is that it should not be started with a load, since all the current required to give a big starting torque must be sent through the armature. The load should be applied gradually when the motor has got up speed.

Armature Reaction

Armature reaction is an interesting phenomenon in motors and dynamos.

If the diagram (Fig. 160) shows the armature of a motor revolving in a counter-clockwise direction, it will be found, by careful application of the left-hand and corkscrew rules, that the armature is magnetized in the direction of the

arrows. The subsidiary poles, which thus appear at top and bottom of the armature, upset the field, which no longer goes straight from N to S.

This may weaken the effective field in a very serious manner. It can be counteracted by putting small subsidiary poles at

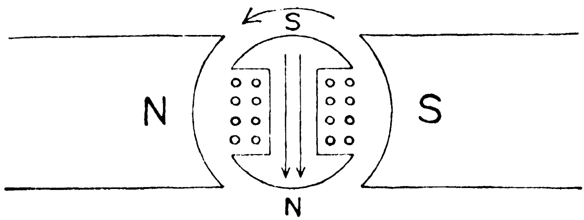


FIG. 160.

right angles to the main poles, as shown in Fig. 161, in order to induce on the armature poles opposite to those due to armature reaction.

In the case of a dynamo revolving in the same direction, the direction of flow of the current will obviously be reversed, and, with it, the direction in which the core is magnetized by armature reaction. So the subsidiary poles must be interchanged.

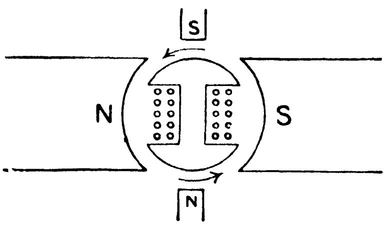


FIG. 161.

Numerical Examples

Problem 5.—A series motor of resistance 2 ohms runs at 2000 R.P.M. when taking 20 amps. at 200 volts under load. What will be its speed if the load is reduced till the current is 10 amps. if the field is proportional to the field-current?

We have for the series motor

$$\text{Back E.M.F.} \propto \text{speed} \times \text{current}.$$

Thus in this case

$$\frac{\text{New speed}}{\text{Old speed}} = \frac{\text{New back E.M.F.}}{\text{New current}} \Big/ \frac{\text{Old back E.M.F.}}{\text{Old current}}$$

$$= \frac{\text{New back E.M.F.} \times \text{old current}}{\text{Old back E.M.F.} \times \text{new current}}.$$

$$\begin{aligned}\text{Now the old back E.M.F.} &= 200 - 2 \times 20 \\ &= 160 \text{ volts.} \\ \text{New back E.M.F.} &= 200 - 2 \times 10 \\ &= 180 \text{ volts.}\end{aligned}$$

$$\therefore \quad \frac{\text{New speed}}{2000} = \frac{180}{160} \times \frac{20}{10}$$

$$\text{New speed} = 4500 \text{ R.P.M.}$$

Problem 6.—A shunt motor drives a shunt dynamo which generates current for laboratory uses at reduced voltage. The motor takes 18 amps. at 200 volts, and the dynamo delivers 110 amps. current externally. The system runs at 1050 R.P.M. The field-resistance of the motor is 100 ohms, and its armature-resistance is 0·25 ohm. The field-current of the dynamo is 2 amps., and its armature-resistance 0·1 ohm. Iron and friction losses in the motor are 152 watts, and in the dynamo 184 watts.

Find (1) Power delivered by motor to shaft in H.P.
(2) Torque in shaft in lb.-feet.
(3) Supply P.D. of dynamo.
(4) Overall efficiency of the system.

1 H.P. is 746 watts, and π may be taken as $\frac{22}{7}$.

The rate of supply of mechanical power by the motor = (back E.M.F. $\times$ armature current) $-$ Losses.

The armature current is $\left(18 - \frac{200}{100}\right)$, or 16 amps. So the back E.M.F. is $200 - 16 \times \frac{1}{4}$, or 196 volts.

So the rate of supply of mechanical power

$$\begin{aligned}&= 196 \times 16 - 152 \text{ watts} \\ &= 3136 - 152 \text{ watts} \\ &= 2984 \text{ watts} \\ &= \frac{2984}{746} \text{ H.P.} \\ &= 4 \text{ H.P.}\end{aligned}$$
. . . . Ans. (1)

A torque of x lb.-feet means that $2\pi x$ foot-lb. of work are done in one revolution.

In our case $4 \times 33{,}000$ foot-lb. of work are done in one minute, during which 1050 revolutions are completed.

Thus torque $= \dfrac{4 \times 33{,}000}{2\pi \times 1050}$

$= \dfrac{4 \times 33{,}000 \times 7}{44 \times 1050}$

$= 20$ lb.-feet . . . . Ans. (2)

The electrical power obtained from the shaft by the dynamo

$= (2984 - 184)$ watts

$= 2800$ watts (since losses in dynamo are 184 watts).

Now electrical power in Dynamo = Armature current × E.M.F.

Armature current = External current + field current

$= (110 + 2)$ amps.

$= 112$ amps.

Thus E.M.F. $= \dfrac{2800}{112}$ volts $= 25$ volts.

But the lost volts are $112 \times 0{\cdot}1$ in the armature,

so supply P.D. $= 25 - 11{\cdot}2$

$= 13{\cdot}8$ volts . . . Ans. (3)

Thus overall efficiency $= \dfrac{13{\cdot}8 \times 110}{200 \times 18} \times 100\%$

$= 42\frac{1}{6}\%$ . . . . Ans. (4)

CHAPTER XVI

A HISTORICAL NOTE

THE history of Electricity and Magnetism, up to about 1850, is the history of a few great men. Since then it has gradually appeared that a knowledge of electricity is the foundation of a knowledge of all material things, and advances have been the result of the cumulative effect of the discoveries of an immense number of workers. Future advances are likely to be made in the same way. The early course of discovery was an amazing one. A close relation between Electricity and Magnetism was suspected for more than 200 years, but none was discovered until Oersted performed his experiment in 1820. In the next fifteen years the foundations of electromagnetism were laid by Faraday, Ampère, Gauss, and Ohm; and more of the fundamental work was done in these years than in the 200 years before or the sixty years after them. Since 1835 three new doors of electrical discovery have been opened, when Maxwell published his *Electromagnetic Theory of Light* in 1864, when the Electron was discovered by J. J. Thomson in 1897, and in recent years when Bohr in 1913 applied Planck's Quantum Theory to account for line-spectra of atoms, and as a result of this in the few years before 1930 when the theory of quantum mechanics appeared to resolve the contradictions of the wave theory and the quantum theory. If we regard the quantum theory and wave-mechanics as represented by the opening of a single door, we may say that save for these three discoveries, all advances have been the logical consequence of former work.

At present, in 1930, a great synthesis of the relativity theory, the quantum theory, and the classical theory of electromagnetism and light is being attempted. It should thus include gravitation and astronomical theory generally, the laws of mechanics, and all atomic and chemical laws. The

result of this may be not only a complete recasting of electrical theory as we now understand it, but also a drastic revision of our view of the whole of experience. Our electrical theory has grown like a ramshackle farmhouse which has been added to, and improved, by the additions of successive tenants to satisfy their momentary needs, and with little regard for the future. We regard it with affection. We have grown used to the leaks in the roof, and are content to speak of the electronic charge as negative. We are content to think of the charge on an electrostatic conductor as continuous rather than atomic, and we do not mind the long and devious passage from dairy to store-room. We are accustomed to deal with three different sets of electrical units, none of them reasonably related to the mechanical units, and we almost admire the miscellaneous and disconnected lean-to sheds which have long disfigured our walls. But our haphazard house cannot survive for ever, and it must ultimately be replaced by a successor whose beauty is of structure rather than of sentiment.

Dr. William Gilbert, Physician to Queen Elizabeth, was the first experimenter in Electricity and Magnetism to record his work in a scientific treatise (1600), *De Magnete Magneticisque corporibus, et de magno magnete Tellure* (On the Magnet and Magnetic Bodies, and on the great Magnet, the Earth). He discovered the repulsion of like poles and the attraction of unlike poles, and encountered the terrestrial magnetic elements, which he understood and (so far as he could) measured. His "terella" or "little earth" gave an illustration of the general form of the earth's magnetic field. He also experimented with the frictional electrostatic charge, and named insulators "electrics" because they could show the phenomenon, and conductors "non-electrics" because (if held in the hand) they could not.

Du Fay produced the "two-fluid" theory of electricity in 1730, and distinguished between "vitreous" (positive) and "resinous" (negative) charge. Mosschenbrock accidentally discovered the Leyden Jar at about the same time. In 1750 Benjamin Franklin introduced the "one-fluid" theory of electricity and (by an unlucky chance) introduced the terms "positive" and "negative" to express vitreous and resinous electricity. He also invented the lightning-conductor.

In about 1765 Priestley, the chemist, demonstrated that the charge on an electrostatic conductor was all on the surface, and, by a brilliant inference, deduced the Inverse Square Law.

The first serious quantitative work ever done in connection with electricity was by the Hon. Henry Cavendish, the third son of the Duke of Devonshire, a strange and solitary man, but a noble one in all senses. Not only was he a great mathematician and theorist, but he was probably the most brilliant experimenter the world has ever seen, and was in every way worthy to give his name to the Cavendish laboratory at Cambridge, the centre of physical research in England, and perhaps in the world.

His marvellous experiments verified Newton's law of gravitation by direct laboratory experiment, and probably led him to suspect the inverse square law for electric charges. He repeated Priestley's experimental and theoretical work in 1771 with far greater accuracy and completeness, and established the foundations of electrical science for all time. He also made condensers, and originated the ideas of potential and capacity, in addition to that of quantity of charge.

At about the same time Galvani accidentally discovered the continuous electric current, but did not recognize its nature; and Volta, with his pile of zinc and copper plates separated by wet flannel, produced a permanent potential difference by non-electrostatic means, and started the idea of the electric cell.

At about this time also, Coulomb made a direct verification of the inverse square law, both for magnetic poles and for electric charges, with his torsion balance, invented by John Michell, a Yorkshire clergyman, and a friend of Priestley and Cavendish.

In 1807 a new era began, when Oersted announced that he would try the effect of an electric current on a magnetic needle.

Though it had always been suspected that magnetism and electricity were connected, no relation between them was ever established till now.

In 1820 Oersted succeeded in his aim—accidentally, at the end of a lecture, as it happened. Though he did no more in the matter, the greatest honour is due to him, for he really started the ball rolling.

Ampère immediately repeated and investigated Oersted's

experiment, and founded on it his mathematical theory of the measurement of current by means of its magnetic field. This theory is explained in Part II, Chapter II.

This work was, and remains, the mathematical foundation of electromagnetic theory. During the same period Gauss was laying the mathematical foundation of the theory of static magnetic and electric fields.

In spite of all this advanced work, the most elementary laws of the behaviour of steady currents were not yet understood. In 1827, as the result of a long series of experiments, G. S. Ohm published the law which bears his name. His work was received with an amazing incredulity, and in consequence of it he lost his teaching post. His conclusions were not universally accepted until some six years after the publication of his paper.

During this period Davy and Grotthus had done preliminary work on Electrolysis, and Seebeck had discovered Thermo-electricity.

After the work of Ampère, Ohm, and Davy, all was ready for Faraday, who in his long life was to prove himself the greatest of all experimental philosophers.

We have already said something rather like this about Cavendish, but really the two men are so different that they cannot be compared. Cavendish was the more brilliant experimenter, because he carried out a few definite and exceedingly difficult experiments with extreme skill, and he was also a very notable mathematician. He was not the far-seeing originator of new ideas that Faraday was, while Faraday was no mathematician, and was not primarily interested in exact measurement, as Cavendish was. Faraday was the greatest of experimental philosophers, and Cavendish the most brilliant experimenter.

Faraday discovered electromagnetic induction in 1831, and formulated its quantitative laws. In 1835 he investigated electrolysis, and formulated its quantitative laws.

He abandoned the unfruitful notion of action at a distance, which has been the hindrance to Newton's theory of gravitation until the time of Einstein, and suggested our modern conception of the electric and magnetic fields. He invented the invaluable conception of lines of force. He forecast the

relation between electromagnetism and light. And in general the ideas originating from him have proved the inspiration of succeeding physicists all over the world.

In spite of his kindliness and charm, he was a wit. To Gladstone's futile question, as to what the use of one of his experiments could be, he replied, " Why, sir, you will soon be able to tax it." And on another occasion he reminded a lady, who asked the same sort of question, of Franklin's answer, " Madam, what is the use of a new-born baby? "

Faraday was the son of a blacksmith and an uneducated farmer's daughter. His parents showed no special signs of intelligence, and he himself had no education worth speaking of. He served a seven-years' apprenticeship to a bookbinder, and became an unskilled laboratory assistant at twenty-two. Kindly, generous, happily married, surrounded by friends and loving children, he lived one of the happiest lives that have been lived, though he had no children of his own.

Two extracts follow, from the notebook in which Faraday made the first records of the experiment of August 29, 1831, when he discovered the induction of one current by the making or breaking of another; and of October 17, when he discovered the production of an induced current in a conductor by the motion of a magnetic field near it.

The trouble of a very careful study of the wording of these extracts will be well repaid.

Each numbered paragraph contains a single idea, experimental or theoretical. Separate deductions from the same experiment appear in different paragraphs, as in Paras. 5 and 58. All relevant facts are put down with great terseness. The dashes and absence of punctuation suggest that parts of sentences were put down immediately as the corresponding part of the experiment was done (*e.g.* Para. 57).

Aug. 29, 1831.

1. Experiments on the production of Electricity from Magnetism, etc., etc.

2. Have had an iron ring made (soft iron), iron round and $\frac{7}{8}$th inches thick and ring 6 inches in external diameter. Wound many coils of copper wire round, one half the coils being separated by twine and calico—there were 3 lengths of

wire each about 24 feet long and they could be connected as one length or used as separate lengths. By trial with a trough each was insulated from the other. Will call this side of the ring A. On the other side but separated by an interval was wound wire in two pieces together amounting to about 60 feet in length, the direction being as with the former coils; this side call B.

3. Charged a battery of 10 pr. plates 4 inches square. Made the coil on B side one coil and connected its extremities by a copper wire passing to a distance and just over a magnetic needle (3 feet from iron ring). Then connected the ends of one of the pieces on A side with battery; immediately a sensible effect on needle. It oscillated and settled at last in original position. On *breaking* connection of A side with Battery, again a disturbance of the needle.

4. Made all the wires on A side one coil and sent current from battery through the whole. Effect on needle much stronger than before.

5. The effect on the needle then but a very small part of that which the wire communicating directly with the battery could produce.

Oct. 17, 1831.

55. Have prepared two pieces of apparatus.

N, a piece of musket barrel $\frac{7}{8}$ of inch in diameter variable in thickness but about $\frac{1}{16}$ of an inch—this was covered by one piece of copper wire 61′ 4″ long in a helix passing from end to end (in the same direction) and back again so as to surround the barrel 4 times.

56. O, a cylinder hollow of paper covered with 8 helices of copper wire going in the same direction and containing the following quantities:

	ft.	in.
1. outermost	32	10
2. ————	31	6
3. ————	30	
4. ————	28	
5. ————	27	
6. ————	25	6
7. ————	23	6
8. ————	22	0
	220	4

put exclusions of projecting ends all separated by twine and calico. The internal diameter of paper cylinder was $\frac{13}{16}$ of inch in diameter—the external diameter of whole 12 inches and the length of copper helix (as a cylinder) $6\frac{1}{2}$ inches.

57. Expts. with O. The 8 ends of the helices at one end of the cylinder were cleaned and fastened together as a bundle. So were the 8 other ends. These compound ends were then connected with the galvanometer by long copper wires—then a cylindrical magnet $\frac{3}{4}$ inch in diameter and $8\frac{1}{2}$ inches in length had one end just inserted into the end of the helix cylinder—then it was quickly thrust in the whole length and *the galvanometer* needle moved—then pulled out and again the *needle moved but* in the opposite direction. This effect was repeated every time the magnet was put in or out and therefore a wave of electricity was so produced from the *mere approximation of a magnet* and not from its formation in situ.

58. The needle did not remain deflected but *returned to its place* each time. The order of motions were inverse as in former expts.—the motions were in the direction consistent with former expts., *i.e.* the indicating needle tended to become parallel with the exciting magnet being on the same side of the wire and poles of the same name in the same direction.

59. When the 8 helices were made one long helix the effect was not so strong on the galvanometer as before, probably not half so strong—so that it is *best* in pieces and combined at the end.

60. When only one of the 8 helices was used it was least powerfull, hardly sensible.

Before Clerk Maxwell's work appeared, Lord Kelvin invented the moving-coil galvanometer and the ballistic galvanometer, and Joule measured the mechanical equivalent of heat, thereby turning Kelvin's attention to the fundamental scale of temperature.

In 1850 Kelvin turned his attention to telegraphy over long distances, and in 1858 the first electric cable was laid. He also made the use of the compass for navigation a really practical proposition.

Maxwell was born in 1831, the year of Faraday's discovery of electromagnetic induction.

When he was an undergraduate, he read Faraday's *Experimental Researches*, and they determined his future.

He noticed that the ratio of the electromagnetic unit of charge to the electrostatic unit of charge has the dimensions of a velocity, and that that velocity is about 3×10^{10} cm. per second, practically the same as the velocity of light. His brilliant conception of the displacement-current in space led to the electromagnetic theory of light, which showed, as a result of ten years of profound mathematical research, that light is simply electromagnetic wave-motion whose wave-lengths cover a certain range.

He thus inferred that electromagnetic waves of other wave-lengths—that is, wireless waves—existed, and predicted their discovery.

This discovery was made by Hertz in 1888, and wireless was made a practical proposition by Marconi and Fleming in researches from 1895 onwards.

In the middle nineties modern physics began with the discovery of the electron by J. J. Thomson, of radioactivity by Becquerel and the Curies, and of *X*-rays by Röntgen. From the turn of the century to the 1914–18 war it continued with the discovery of the Quantum Theory by Planck and the Relativity Theory by Einstein. These theories, applied to Rutherford's conception of the nuclear atom, led us by 1945 to the beginning of the era of available atomic energy.

We take the electron as a commonplace now, and forget how shocking it was for physicists to be expected to believe in atoms of electricity, in spite of Faraday's laws of electrolysis. Faraday dislikes the idea (*Experimental Researches*, 1833, sections 852, 869), and Maxwell (*Electricity and Magnetism*, published 1873, 3rd edition, pp. 380, 381), after speaking of a molecule of electricity as attached to an ion in electrolysis, adds: "This phrase, gross as it is, and out of harmony with the rest of this treatise, will enable us at least to state clearly what is known about electricity."

Fearing that so horrible a theory may be taken seriously, he adds on the next page: "It is extremely improbable that when we come to understand the true nature of electrolysis we shall retain in any form the theory of molecular charges, for then we shall have secured a sure basis, upon which to form a

true theory of electrical currents, and so become independent of these provisional hypotheses."

Dr. Johnstone Stoney was the first to state clearly a belief in the atomic nature of electricity, at a British Association meeting in 1874. He made the first estimate of its size (based on the amount of electricity required to liberate one gramme of hydrogen in electrolysis and the estimate from kinetic theory of the number of atoms in a gramme-molecule) and in 1891 he first coined the term " electron."

Helmholtz, in his Faraday lecture at the Royal Institution in 1881 committed himself definitely to the atomicity of electricity in electrolytic solutions, though not in metals. It seems that in this matter the most penetrating guessers were Faraday and Stoney.

The discovery of the electron is generally said to have been made by J. J. Thomson in the Cavendish Laboratory in 1897. Strictly, what he then found was the value of $\frac{e}{m}$ for the particles (*i.e.* electrons) in the cathode rays of a discharge-tube, and so to measure the constants of an electron detached from matter; in all previous measurements the electron had been attached to some sort of ion, either electrolytic or gaseous. In such measurements only the charge was measured, and the mass was ignored. In the same year, 1897, $\frac{e}{m}$ for electrons was also measured independently by Kaufmann and by Wiechert.

The first direct measurement of e for gaseous ions was made by Townsend, also in 1897, using C. T. R. Wilson's expanding cloud method, and it was followed by Thomson, using a similar method. A great improvement on this method was made by H. A. Wilson in 1903, and Lenard was in 1905 awarded the Nobel prize for showing that cathode rays could pass right through aluminium foil, and could not therefore be atoms of any ordinary kind. He also showed by these experiments that atoms themselves were not solid, but were empty enough for cathode-ray particles to pass right through them.

R. A. Millikan in America made another great advance on H. A. Wilson's method of measuring e. He used an oil-cloud instead of a water-cloud, and in a long series of determinations

evolved what still (1948) remains the standard direct way of measuring e. The accepted value comes indirectly from Avogadro's number N and the Faraday .

Positive rays, originally discovered by Goldstein in 1886, were effectively investigated by Thomson's parabola method in 1906, and with much greater accuracy by F. W. Aston with his mass-spectrograph, also at Cambridge, from 1919 onwards. Thomson discovered the isotopes of neon in 1912, and Aston an enormous range of isotopes from 1919 onwards.

Following the preliminary work of Rutherford and the Curies on radioactivity, Crookes in 1903 discovered the scintillations due to the α-particles from radium, and so led Rutherford to the discovery of the α-particles themselves, later the main tools of his work on atomic structure.

In 1906 Max Planck of Berlin (he died in 1947) proposed the quantum theory to account for the peculiar distribution of energy in the spectrum of a radiating black body; and in 1912, by an amazingly bold leap of speculation, Niels Bohr applied Planck's quantum theory to Rutherford's conception of the hydrogen atom as a minute heavy nucleus with a single light electron describing an orbit about it at a great distance, and produced a complete and exact solution for the extremely complex hydrogen spectrum.

The two last great pieces of research before the 1914–18 war were both about X-rays, which had been accidentally discovered by Röntgen in 1895 through the fogging of a covered photographic plate left near a discharge-tube, and had been shown to be like light, but of much shorter wave-length.

The young Oxford physicist, H. G.-J. Moseley, proved the existence of a simple relationship between the characteristic X-rays of an element and its atomic number, and for the first time gave us a true understanding of the periodic table and an exact meaning for the atomic number. Moseley was killed at Gallipoli in August 1915 at the age of 26. At about the same time, W. H. and W. L. Bragg, father and son, successfully used X-rays to investigate the structure of crystals, and so opened up a way to the sub-microscopic investigation of the structure of all matter.

The story of physics from 1918 onwards is continued in Chapter I of Part III

CHAPTERS I–IV. ELECTROSTATICS

General Questions

1. How has the inverse square law of force been verified for electric charges? How does this law provide a basis for defining the unit of charge?

2. What is the evidence that the charge which flows from a voltaic cell is of the same kind as that produced by a Wimshurst machine?

3. What is meant by surface-density of electrification? How does it vary on the surface of an irregularly shaped insulated conductor? How could you illustrate your statement experimentally?

4. Describe some form of electrical induction machine. Explain the discharging action of points and flames and the phenomenon of the electric wind.

5 Explain the process of charging by induction, with special reference to the electroscope and the electrophorus.

6. Give an account of Faraday's contributions to electrostatics.

7. Explain what is meant by the statement that one conductor is at a higher electrical potential than another. How is this difference of potential defined quantitatively?

8. What do you mean by the capacity of a conductor? On what factors does it depend?

9. Explain what is meant by a condenser, and deduce an expression for the capacity of a parallel-plate condenser. How would you compare the capacities of two condensers?

10. What is meant by the electric intensity at a point?

Explain shortly the relations between charge, intensity, potential, and capacity.

11. An insulated uncharged sphere is placed between a positively charged plate and the earth's surface. Discuss the distribution of potential of the system.

12. Explain the principle of the Guard-Ring condenser. How has it been adapted to give an absolute measure of potential?

Deduce an expression for the potential measured.

13. What is meant by the dielectric constant of a medium?

How does it affect (*a*) the capacity of condensers, (*b*) the dissociation in electrolytes?

14. Discuss the changes of charge, potential, and capacity when a sheet of glass is put between the plates of a condenser (*a*) when the condenser is charged and the charged plate insulated,

(*b*) when the plates of the condenser are connected to a source of constant potential difference.

15. How and why is the situation affected when a Leyden jar which is being charged by a Wimshurst machine is placed on a large block of paraffin wax instead of on the table?

16. Condensers of capacities C_1, C_2, and C_3 are connected (*a*) in parallel, (*b*) in series. Find expressions for the capacities of the compound condenser so formed in each case

17. Find an expression for the energy of a charged conductor or condenser. Can this energy be said to be located in any special place?

How would you calculate the heat evolved when a condenser is discharged?

18. From considerations of energy deduce the direction of the force on the glass sheet in both cases in Question 14.

19. Obtain an expression for the heat evolved when two condensers at different potentials are connected in parallel.

20. Describe a gold-leaf electroscope. What does it measure, and what factors determine the position in which the leaf comes to rest?

21. Describe carefully Faraday's Ice-pail experiment, and show what may be deduced from it? How do you account theoretically for the behaviour observed?

22. How would you (*a*) compare the magnitudes of unknown charges on insulators of different dimensions, (*b*) demonstrate that the charges produced by friction are equal and opposite?

23. How would you use an electroscope (*a*) to find the sign of an unknown charge of an insulated conductor, (*b*) to find which pole of a cell is positive?

24. Why do you obtain no visible deflection when an electroscope is connected directly to a Leclanché cell, and how could you succeed in producing a visible deflection with the help of the cell?

25. Discuss the conception of lines of force. Do you regard them as real? If you do, explain how to demonstrate that they are real. If you do not, explain how they can be useful.

26. Enumerate the properties of lines of force, explaining which are assumed and which deduced, and which are quantitative and which qualitative.

27. Use the conception of lines of force to show that the potential of an insulated charged spherical conductor in free space is the same as the potential of a point at a distance equal to the radius from a point-charge equal to the total charge on the sphere. What is this potential? Deduce the capacity of the sphere.

28. Describe the construction of the quadrant electrometer. How may it be used (*a*) to compare the E.M.F.s of two cells; (*b*) to compare the charges on two conductors; (*c*) to compare the capacities of two condensers of large capacity of the order of 1 microfarad; (*d*) to compare an alternating E.M.F. with a direct E.M.F.?

29. Describe a modern sensitive form of electroscope suitable for quantitive measurements, explaining clearly what it is that the electroscope measures.

30. Describe the preliminary adjustments you would make when setting up a quadrant electrometer. What troubles would you expect, and what precautions would you take? Give reasons for your answers.

Numerical Questions

31. Find the strength of two equal charges which exert a force of 4 dynes on each other when 6 cm. apart *in vacuo*.

32. Find the strength of two equal charges which exert a force of 1 gm. weight on each other when 2 cm. apart *in vacuo*.

33. What force does a charge of 24 e.s. units exert on a charge of 16 e.s. units 4 cm. away from it?

34. Two small equal pith-balls hang by light non-conducting threads from the same point. They are charged, and fly apart so that the angle between the strings is 30°. The mass of each ball is 1 gm. The strings are 10 cm. long. Find the charge on each ball.

35. Equal charges of Q units are at the corners of a square of side *a*. Find the direction and magnitude of the intensity at the mid-point of one of the sides, and the magnitude and direction of the force on each charge.

36. Charges of + Q and − Q are at the ends of the base of an isosceles triangle whose height is twice its base-length. Find the direction and magnitude of the electric intensity at the third vertex. The length of the base is $2a$. (The broadside position of a bar-magnet will be found to provide a useful analogy.)

37. Find the intensity and potential 10 cm. from a charge of 20 e.s. units *in vacuo*.

38. Point charges of + 100, − 100, + 200 are at the vertices of an equilateral triangle of side 8 cm. Find the force on a charge − 2 at the centre of the triangle, and at a point midway between the equal charges. How much work would be done in taking this charge from one point to the other?

39. Find the potential gradient in volts cm. at a distance of 5 cm. from a charge of 4 e.s. units *in vacuo*.

40. A surface has a density of charge of one-thousandth of a micro-coulomb per sq. cm. Find the potential gradient in volts/cm. just outside it.

41. The electric field just outside the earth's surface is about + 3 volt/cm. Find the surface density of the earth's charge, and show that it is negative.

42. The earth's radius is $6{\cdot}4 \times 10^8$ cm.

From the result of the last question, find the total charge on the earth's surface, and the potential in volts at which it would be if it were an isolated conductor in free space.

43. If the greatest electric intensity the air can support is 30,000 volts per cm., find the maximum charge in coulombs which could be held by a sphere of 10 cm. radius, and the largest possible surface density of charge.

44. What is the radius of a sphere whose capacity is 1 microfarad?

45. What is the capacity, in microfarads, of a Leydon jar, 20 cm. high of diameter 10 cm., if the thickness of the glass is 1 mm. and its dielectric constant is 6?

46. How much heat is evolved when a spark 2 cm. long can just pass across the Leyden jar of Question 45? (See Question 43 for the potential.)

47. A slab of glass 2 cm. thick of dielectric constant 6 is introduced between the plates of a parallel plate condenser 5 cm. apart. In what ratio is the capacity increased?

48. Condensers of capacity 2, 3, and 4 microfarads are charged to potentials of 300, 200, and 100 volts respectively. They are then connected in parallel. How much energy in ergs is evolved? What would their combined capacity be if they were arranged in series?

49. A drop of oil of radius 2×10^{-4} cm. and density 0·8 gm./c.c. is held suspended between two parallel plates 0·2 cm. apart. The charge e on the drop is 1 electron. What is the potential difference in volts between the plates?

$$g = 981 \text{ cm./sec.}^2 \quad e = 4{\cdot}774 \times 10^{-10} \text{ e.s.u.}$$

50. A sphere of radius 20 cm. is charged with 10 e.s. units, and deflects an electrometer 40 divisions. It is then connected by a long wire to a 10 cm. sphere and deflects the electrometer 30 divisions. What is the capacity of the electrometer?

51. Charges of + 10 and − 20 units on equal conducting spheres are brought together so that they share their charge,

and separated again to the same distance as before. Find the ratio of the present force between them to the original force.

52. Two parallel plate condensers of area 10 and 20 sq. cm. respectively, with a distance of 1 mm. between their plates, are charged so that they each have a force of 1 gm. weight between the plates. They are connected in parallel. How much heat is dissipated?

53. Two spheres of diameters 10 and 16 cm. are given charges of 4 and 3 e.s. units respectively. They are then connected by a long thin wire. How much energy is dissipated from the wire?

54. An electron of charge $4{\cdot}774 \times 10^{-10}$ e.s. units and mass $8{\cdot}8 \times 10^{-28}$ gm. is released with zero velocity from the negative plate of a parallel plate condenser whose plates are 2 cm. apart *in vacuo*. The charge on the condenser is 3 e.s. units per sq. cm. What is the electron's acceleration, and how long does it take to get to the positive plate?

55. A condenser of 1 microfarad capacity is in series with another of 2 microfarads. The outer plate of the latter is earthed, and a potential of 3000 volts is applied to the outer plate of the former. Find the potential of the inner plate and the charge on each plate.

56. The moving plate of a Kelvin electrometer is moved 0·4 cm. between its two adjustments. The mass used is 0·2 gm., and the area of the plates is 50 sq. cm. Find in volts the measured potential difference.

57. Find the mass required in a Kelvin electrometer to give force equal to the force of attraction between the plates when they are 0·4 cm. apart, of area 20 sq. cm., and at a potential difference of 1200 volts.

58. A parallel plate condenser is insulated at a potential difference of 100 volts. A plate of glass is dielectric constant 6 and thickness 1·0 cm. is slid between the plates. How much further apart must the plates be moved to bring the potential difference to its original value?

59. A parallel plate condenser is at a potential of 1000 volts when its charge is 0·01 coulomb. How much work must be done to double the distance between the plates (*a*) when the charged plate is insulated, (*b*) when it is kept at 1000 volts?

60. Two raindrops a long way apart have radii of 2 and 3 mm. respectively. Their potentials are 50 and 60 e.s. units respectively. What will be the change in energy of charge if they coalesce, and what will be their potential?

CHAPTERS V–VII. MAGNETISM

General Questions

61. Describe an experimental method of verifying the law of inverse squares for magnetic poles.

62. Explain what is meant by the magnetic moment of a magnet, and discuss the advantage gained by introducing this conception.

63. What do you mean by a magnetic field? What factors are necessary to define it? How would you determine the strength and direction of the earth's magnetic field at any point?

64. What do you mean by magnetic pole? Wherein does it differ from electric charge? Define unit magnetic pole.

65. What is meant by "dip"? Describe the dip circle and explain how it is used. Describe and explain the behaviour of the needle of a dip circle during a complete turn of the instrument about a horizontal axis (*a*) in England, (*b*) at the magnetic equator.

66. Discuss briefly the possible causes of the earth's terrestrial magnetism.

67. Discuss the variations in the earth's magnetic field, and suggest reasons for them.

68. How does the period of an oscillating magnet depend on the strength of the field in which it is placed?

69. Describe how you could determine from first principles the magnetic moment of a magnet and the strength of the horizontal component of the earth's magnetic field.

70. Deduce expressions for the intensity due to a bar magnet at points on its axis and on its perpendicular bisector. Explain how these expressions can be modified if the magnet is short, and how they can be used to verify the inverse square law.

71. Explain the neutral point method of finding the magnetic moment of a bar magnet. Sketch the lines of force due to a magnet (*a*) with its N. pole pointing N., (*b*) with its N. pole pointing S., (*c*) with its N. pole pointing E. in the earth's field.

72. Explain how to compare the magnetic moments of two magnets by means of a vibration magnetometer, with particular attention to the elimination of the earth's magnetic field from the calculations.

73. What are the resemblances and differences between the earth's magnetic field and that which would exist if the earth were an iron sphere uniformly magnetized parallel to its axis?

NUMERICAL QUESTIONS

74. Find the force exerted by a pole of strength 20 webers on a pole of strength 10 webers if they are 5 cm. apart *in vacuo*, and the intensity due to them at a point 5 cm. from each.

75. A bar-magnet has pole-strength 100 webers, and magnetic moment 2000 units. Find the magnetic intensity due to it at a point 20 cm. from its centre (*a*) when this point is on the axis, (*b*) when the point is on the perpendicular bisector, (*c*) when the point and one pole form the ends of the base of a right-angled isosceles triangle with the other pole at the third vertex.

76. Two short bar-magnets of magnetic moments 300 and 400 units respectively are fixed with their axes horizontal and their centres in the vertical line about which the system is free to oscillate. Compare the times of oscillation in the earth's field (*a*) when their axes are in the same direction, (*b*) when their axes are in opposite directions, (*c*) when their axes are at right angles.

77. A magnet 10 cm. long has a neutral point, when its S. pole points N., at a distance of 30 cm. from its centre. The total intensity of the earth's field is 0·47 oersted, and the dip is 67°. Find the magnetic moment and pole-strength.

78. A dip needle removed from its bearings and suspended so as to swing horizontally makes a complete vibration in 10″. Its moment of inertia is 800 units. When it is on its bearings a mass of 0·1 gm. at a distance of 8 cm. from its centre makes it come to rest in a horizontal position. Find the dip.

79. The apparent dip is 60° when the plane of the dip circle is pointing N.N.W., and 55° when it is pointing E.N.E. Find the dip. [Directions Magnetic, not Geographical.]

If the apparent dip in a particular direction were observed at the same place as 62°, what would it be observed as, in a direction perpendicular to this?

80. A short bar-magnet oscillates in 14 seconds in the earth's field, and causes a deflection of 45° in a tangent galvanometer placed at 30 cm. from its centre in a magnetic westerly direction on the line of its axis. The moment of inertia of the magnet is 2160 units. Find its magnetic moment and the strength of the horizontal component of the earth's magnetic field.

81. A vibration magnetometer oscillates in 10″ in the earth's field alone, and in 8″ when it is on the axis of a magnet pointing N. How long will it take to oscillate if the magnet is reversed?

82. Two bar-magnets of pole-strengths 100 webers and mag-

netic moments 2000 units lie parallel 5 cm. apart. Find the force between them (*a*) when they point in the same direction, (*b*) when they point in opposite directions.

CHAPTERS VIII, IX. STEADY CURRENTS AND OHM'S LAW

General Questions

83. State and explain Ohm's Law, and use it to explain the principles of the Wheatstone Bridge and the Potentiometer.

84. Employ Ohm's Law to deduce from first principles expressions for the resistance of two conductors of known resistance (*a*) in series, (*b*) in parallel

85. How would you verify Ohm's Law directly (*a*) ideally, (*b*) with elementary apparatus?

86. Describe the construction of some common form of voltmeter and explain its action.

87. Describe the construction of some common form of ammeter and explain its action.

88. Describe the Post-Office Box. What special advantages (if any) has this particular arrangement of the Wheatstone Bridge?

89. Describe the tangent galvanometer, and obtain an expression relating current and deflection for it. What is it useful for?

90. Describe the Carey-Foster Bridge. What is it useful for?

91. What methods would you use for the measurement of (*a*) very high, (*b*) very low resistance?

92. Define the ampere, the volt, and the ohm, and explain how they are related to some absolute system of units.

93. State Kirchoff's Laws for the distribution of currents in a network of conductors, and explain how they may be deduced from elementary principles.

94. Explain what is meant by the E.M.F. of a cell. Why is it different from the potential difference between the terminals when the cell is supplying current? What can be deduced from observations of the E.M.F. and P.D.?

95. Explain the principle of the Astatic Galvanometer.

96. What are the principal effects of the passage of an electric current? Which of them can, and which cannot, be made the basis of a method of measuring current? Why are some favoured in this respect?

97. Obtain an equation relating the resistance of a length of wire to the specific resistance, and define the latter term.

98. Define electrical resistance, and show how the definition depends on Ohm's Law. Suggest cases which occur to you in which Ohm's Law is not obeyed. What meaning do you attach to the term "resistance" in such cases?

Numerical Questions

99. A dynamo supplies a current of 12 amps.

Its driving engine works at the rate of 2·22 h.p. The System is 80% efficient. Find the E.M.F. of the dynamo (1 joule = 0·74 ft.-lb.).

100. A 200-volt motor raises water at the rate of 1200 lb. a minute from a well 50 feet deep. The efficiency is 60%. What current is taken?

101. A lamp is marked 50 watt 200 volt. What current does it take, and what is its resistance on a 200-volt circuit?

102. A 200-h.p. engine drives a dynamo which supplies 25 amps. at 5000 volts. What is the efficiency of the system?

103. A galvanometer of resistance 12 ohms gives maximum deflection for a current of 0·02 amp. What is the greatest P.D. it will measure, if used as a voltmeter, (*a*) as it stands, (*b*) with a series resistance of 188 ohms?

104. A galvanometer has a deflection of 1 microamp. per scale division, and is to be used as a voltmeter reading 1 millivolt per scale division. Its resistance is 100 ohms. What series resistance should be added?

105. A dynamo delivers 200 amps. to mains at a P.D. of 280 volts. The resistance of the leads between dynamo and mains is 0·1 ohm. What is the E.M.F. of the dynamo, and what percentage of the power is lost in transmission?

106. A dynamo generates at 200 volts, and supplies 500 amps. along leads whose total resistance is 0·04 ohm. 300 amps. are taken off half-way along these leads, and the remaining 200 amps. at the end of them. At what P.D. is each part of the current supplied, and what percentage of the total energy generated is usefully employed?

107. Resistances of 2 ohms and 3 ohms are connected in parallel, and carry, between them, 10 amps. If a third resistance is connected in parallel so that the 3-ohm coil only carries 2 amps., the total current remaining the same, what is the value of this third resistance?

108. Explain how the galvanometer in Question 103 could be converted into an ammeter reading to 5 amps.

109. Explain how the galvanometer in Question 104 could be converted into a milliammeter reading 1 milliamp. per scale division.

110. Manganin has a specific resistance of 43×10^{-6} ohms per cm. cube. No. 22 gauge has a diameter of $7{\cdot}11 \times 10^{-2}$ cms. What length of such wire would be necessary in Questions 108 and 109?

111. A galvanometer gives full-scale deflection with 0·002 amp. Its resistance is 100 ohms. An external resistance is added to make it into a voltmeter reading to 1 volt. The galvanometer is now shunted with a 10-ohm coil. What voltage is now required to give full-scale deflection?

112. The specific resistance of German-silver is 22×10^{-6} ohms/cm.-cube. What is it in microhms/inch-cube?

113. A wire has a resistance of 1·2 ohms per metre and a specific resistance of $47{\cdot}5 \times 10^{-6}$ ohms/cm. cube. What is its diameter?

114. If the P.D. across the terminals of a battery is 4·5 volts on open circuit, and 4·0 volts with an external resistance of 8 ohms, find the E.M.F. and internal resistance of the battery.

115. A cell with an internal resistance of 1 ohm gives a current of 0·5 amp. when a 2-ohm resistance is connected across its terminals. Calculate its E.M.F. and the P.D. across its terminals.

116. A battery of cells of E.M.F. 9 volts sends 3 amps. through a 1-ohm coil. What current would it send through a 2-ohm coil?

117. A cell has E.M.F. 1·5 volts. The voltmeter reading falls to 1·0 volt when a 2-ohm resistance connects the terminals of the cell. Find the current supplied and the internal resistance.

118. A voltmeter of resistance 2000 ohms is used to measure the E.M.F. of a cell of resistance 100 ohms. Find the percentage error of the voltmeter reading.

119. A cell shows an E.M.F. of 1·09 volts by potentiometer, and 1·06 volts on a voltmeter of resistance 1000 ohms. What is the internal resistance of the cell?

120. A battery of N cells is arranged so that it has $\frac{N}{p}$ lots in series of p sets in parallel.

Show that the current through an external resistance R is $\frac{pNE}{p^2R + NB}$ where E and B are the E.M.F. and internal resistance of each cell.

121. It can be shown (see Appendix 3) that the current supplied in Question 120 is a maximum when $p = \sqrt{\frac{NB}{R}}$.

Suppose you have 60 cells of resistance 1 ohm and E.M.F. 1·5 volts, and find the maximum current they can supply to external resistances of 0·01 ohm, 0·1 ohm, 1 ohm, 4 ohms, 15 ohms, 100 ohms.

122. Twelve equal wires, each of resistance 1 ohm, lie along the edges of a non-conducting cube. Find the resistances between the ends of a main diagonal of the cube and the current in each wire if there is a potential difference of 1 volt between these points.

123. The wire AB along one edge of the cube in Question 122 is removed. What is now the resistance between A and B, and the current flowing in each wire for a P.D. of 1 volt between A and B?

124. AB are the battery terminals of a Wheatstone Bridge, and CD the galvanometer terminals. The resistances are as follows:—AC = 1 ohm, CB = 5 ohms, AD = 2 ohms, DB = 4 ohms, CD = 3 ohms. Find the current in CD if the P.D. between A and B is 31 volts, and the resistance from A to B.

125. Cells of E.M.F. 2 volts and 1·5 volts respectively are connected in parallel, and their internal resistances are 0·5 ohm and 1 ohm respectively. They send an external current through a resistance of (*a*) 0·5 ohm, (*b*) 2 ohms, (*c*) 10 ohms. Find the current supplied by each cell, and the current through the external resistance in each case.

126. Repeat Question 125, with the alteration that the 2-volt cell has resistance 1 ohm, and the 1·5-volts cell has resistance 0·5 ohm.

127. The positive poles A and C of two cells are joined by a resistance of 2 ohms, and their negative poles B and D by a resistance of 3 ohms. The mid-point of AC is earthed. At what potential is the mid-point of BD?

AB has E.M.F. 2 volts and resistance 1·5 ohms, CD has E.M.F. 1·5 volts and resistance 0·5 ohm.

128. An accumulator of E.M.F. 40 volts and charging rate 4 amps. is being charged from 100-volt D.C. mains, by means of lamps rated 100-volt 50 watt. How many lamps will be required and how should they be arranged?

129. An H.T. accumulator of E.M.F. 60 volts has a charging rate of $\frac{1}{10}$ amp., and is being charged from the switchboard of Question 128.

How many lamps are required, and how should they be arranged?

130. Four batteries of E.M.F.s E_1, E_2, E_3, E_4 have resistances B_1, B_2, B_3, B_4. Their positive poles are connected together with thick copper wire of negligible resistances. So are their negative poles. What current is supplied by the cell E_1, and what is the common potential difference?

N.B.—Tangent galvanometer questions will be found among questions on electromagnetism.

CHAPTER X. HEATING EFFECTS

General Questions

131. Deduce expressions for the heat developed in a wire in t seconds when a current is flowing through it.

132. Describe Callendar and Barnes' Constant Flow method for determining the mechanical equivalent of heat, deducing the final equation.

133. An electric heater contains two coils of different resistances. Deduce expressions for the rate of production of heat at a given voltage when the coils are used (*a*) separately, (*b*) in series, (*c*) in parallel.

134. Give a general account of the variation of resistance of metals with temperature. How do pure metals differ from alloys in this respect?

How may this variation be used to measure temperature?

135. Give a general account of the Thermo-Electric effect, and explain how it is used (*a*) to measure temperature, (*b*) to measure current.

136. How would you measure the E.M.F. of a thermocouple?

137. Explain shortly the principles and uses of the Thermopile and the Bolometer.

138. What factors determine the steady temperature of a wire carrying a current?

139. Show that the squares of the currents which keep two wires of the same material, but different diameters, at the same temperature are proportional to the cubes of the diameter.

140. Find an expression for the number of calories radiated per second per sq. cm. of surface from a wire carrying a current, in terms of the current, the specific resistance, and the radius of the wire.

Numerical Questions

141. An electric kettle has a resistance of 50 ohms, and is supplied at 200 volts. If 25% of the heat supplied is lost by radiation, find how long it will take to boil water in a copper kettle holding 2 litres, starting from a temperature of 10° C. Mass of kettle = 1000 gms. Specific heat of copper = 0·1.

142. If the kettle of Question 141 had been required to boil the water in 10 minutes under the same conditions, what would its resistance have had to be reduced to?

143. Two wires contain the same mass of metal of the same kind, but one wire is twice as long as the other. Compare their rates of production of heat (*a*) at the same voltage, (*b*) carrying the same current

144. A village requires 100 amps. at 200 volts, which is supplied by a dynamo generating at 210 volts. Find the resistance of the supply mains and the rate of loss of heat from them.

145. A wire of 2 ohms resistance is connected across a cell of E.M.F. 1·5 volts and resistance 1 ohm. Compare the heat generated in the wire with that generated in a wire of 1 ohm resistance in the same position.

146. In Question 145, what value would the external resistance require so that halving it would leave the rate of production of heat the same? What would this rate of production be in watts?

147. A coil whose resistance is 100 ohms at 15° C. increases its resistance by 0·5 ohm per 1° C. rise above that temperature. It loses heat at the rate of 20 watts per degree rise above 15° C. If a P.D. of 200 volts is applied to it, and the air temperature is 15° C., what will be its final steady temperature?

148. A wire whose resistance is directly proportional to the absolute temperature is kept at a temperature of 1000° Absolute by a P.D. of 200 volts. If the rate of radiation of heat is proportional to the 4th power of the absolute temperature, find the P.D. necessary to maintain a temperature of 2000° Absolute.

149. The filaments of a 40-watt 110-volt lamp and a 20-watt 220-volt lamp are run at the same temperature and made of the same material. In what ratio are the filament diameters?

150. In an electric lamp the resistance varies as the absolute temperature, the rate of heat loss as the cube of the absolute temperature, and the candle-power as the 5th power of the absolute temperature. If the candle-power is 24 with a P.D. of 225 volts, what is it with a P.D. of 256 volts?

151. An oil bath is heated by a coil of wire attached to 200-

volt mains, the current being regulated by a series rheostat. A current of 1 amp. will keep the bath at 40° C. when the room-temperature is 15° C. How much resistance must be cut out of the rheostat to maintain the same temperature in the bath if the room-temperature falls to 10° C.?

152. Compare the heat generated in the four arms of a balanced Wheatstone Bridge of resistance 100, 10, 300, 30 ohms respectively. Current enters at junction of 100 and 300 ohm arms.

153. A wire of specific resistance 22×10^{-6} ohms/cm. cube, and diameter 1 mm., is just melted by a current of 10 amps., the air temp. being 15° C., and the melting-point 265° C. Find the emissivity of the surface of the wire (the calories given out per sq. cm. per sec. per 1° C. temp. difference).

CHAPTERS XI, XII. CELLS AND ELECTROLYSIS

General Questions

154. Write a short essay on Electrolysis, indicating how you would verify the laws.

155. State Faraday's Laws of Electrolysis, and give a general account of the Ionic Theory.

156. Describe the effects produced by the passage of an electric current between platinum electrodes through (*a*) a solution of caustic soda in water, (*b*) a solution of copper sulphate.

157. Give an account of what happens when a P.D. of 2 volts is given to a pair of copper plates dipping into (*a*) dilute sulphuric acid, (*b*) copper sulphate solution.

Define the terms anode, cathode, electrochemical equivalent, chemical equivalent, ion.

158. Compare the effects produced by the passage of a current through an electrolytic conductor with the effects produced by the passage of a current through a metal.

159. What is the evidence that a salt is already dissociated into its ions when it is dissolved in water, and is not dissociated by the passage of the current? How do you account for the dissociation?

160. Why does the density of the acid in an accumulator change during charge and discharge? How far can the density of the acid be taken as an indication of the state of discharge?

161. Describe the construction and method of action of the lead-accumulator. What factors determine (*a*) the E.M.F., (*b*) the ampere-hour capacity, (*c*) the watt-hour capacity?

162. Distinguish between the ampere-hour efficiency and the watt-hour efficiency, of an accumulator. What values would you regard as satisfactory?

163. Describe the Daniell cell, and give an account of the processes which take place when it is supplying a current. Is it reversible? If so, why can it not be used as an accumulator?

164. Describe the Leclanché cell, and give an account of the processes which take place when it is supplying a current. Is it reversible? For what is it useful, and why?

165. Describe the construction of an ordinary dry cell.

166. Describe the simple cell formed of zinc and copper plates in sulphuric acid. Why is it useless for practical purposes? Give as many reasons as you can.

167. Give an account of the methods available for determining the internal resistance of a primary cell.

168. Explain the phenomena of local action and polarization. How may they be overcome?

169. What are the sources of the energy of a primary cell? Why does the simple method of calculating the E.M.F. generally fail?

170. How would you find the arrangement of a set of cells which would give the biggest current possible through a given resistance? (See Appendix 3.)

171. Describe the ordinary method of charging an accumulator, showing how to calculate the number of carbon lamps required.

Numerical Questions

172. A potential difference of 2 volts is applied to a water voltameter, and the current is 0·05 amp. The P.D. is then increased to 4 volts, and the current rises to 0·30 amp. What do you deduce from this?

173. A Daniell cell of resistance 0·4 ohm and E.M.F. 1·1 volts is in series with a resistance of 1·8 ohms. How much zinc will be consumed in an hour?

e.c.e. of hydrogen $= 1{\cdot}045 \times 10^{-5}$ gm./coulomb.
Atomic weight of zinc $= 65{\cdot}4$.
Valency of zinc $= 2$.

174. A current which deposits 0·85 gm. of copper in 25 minutes passes through a 2-ohm resistance. A voltmeter gives the potential difference across this resistance as 3·50 volts. What is its error? (e.c.e. of copper $= 3{\cdot}29 \times 10^{-4}$ gm./coulomb.)

175. What mass of hydrogen would be liberated in an hour by

a current of 100 amperes? How many litres at N.T.P. would this occupy, the standard density being 9×10^{-5} gm./c.c.?

176. Find the number of kilowatt-hours required to deposit 2 kg. of copper at a P.D. of 20 volts.

177. The atomic weight of lead is 207·2 and its valency is 2. The atomic weight of silver is 107·88 and its valency is 1. The e.c.e. of silver is 0·001118 gm./coulomb. Lead combining to form $PbSO_4$ gives up 380 cals./gm. The corresponding weight of PbO_2 combining to form $PbSO_4$ gives up 80 cals. What approximate E.M.F. for a storage cell would you expect from these facts by Kelvin's method, p. 211?

178. If the calorific value of hydrogen is $3 \cdot 45 \times 10^4$ cals./gm., find the energy liberated in the electrolysis of water when 1 coulomb passes, and the smallest p.d. required to electrolyze water.

179. The charging rate of an accumulator is 1·5 amps. A number of 220-volt 110-watt lamps are used for charging. How many of these would be used with 220-volt mains?

180. Four cells are charged from 210-volt mains. Each has an E.M.F. of 2·5 volts on charge. The charging current is 0·5 amp. Specify in volts and watts the best lamp to use.

181. A motor takes 0·8 amp. at a P.D. of 10 volts. How many Daniell cells in series, each with E.M.F. 1·1 volts and resistance 0·5 ohm, will be required to run it?

182. Wires of resistance 2 and 3 ohms are in parallel, and a cell of resistance 0·5 ohm and E.M.F. 1·5 volts sends current through them. Find the current through the 2-ohm coil? What is it when the 3-ohm coil is disconnected?

183. A battery of E.M.F. 10 volts and resistance 1 ohm sends current through a circuit ABC. AB is a wire of resistance 7 ohms. BC is a combination of three wires, resistances 4, 5, 6 ohms, in parallel. Find the P.D. across the battery and the current in each wire.

CHAPTERS XIII, XIV, XV. ELECTROMAGNETISM, ELECTROMAGNETIC INDUCTION, DYNAMOS, AND MOTORS

General Questions

184. Give an account of the magnetic effects due to a current. Explain how they are used to define the absolute electromagnetic unit of current.

185. Describe the Tangent Galvanometer, and deduce an expression relating the deflection to the current in amperes.

186. Sketch the lines of force in the neighbourhood of a current flowing in (*a*) a long straight wire, (*b*) a short solenoid, (*c*) a long solenoid, (*d*) a circle.

187. Explain how you would deduce the direction of the magnetic field due to a current, and the direction of the force on a current in a magnetic field.

188. How would you illustrate experimentally the nature and direction of the forces exerted (*a*) by a current on a magnet, (*b*) by a magnet on a current, (*c*) by a current on a current? State the rules summarizing these actions.

189. Describe some form of suspended coil galvanometer and explain its action. On what factors does the couple on a coil in a magnetic field depend?

190. State the laws of electromagnetic induction, and describe experiments illustrating them.

191. State Lenz's Law and illustrate it by explaining the phenomena observed (*a*) when a thick plate of copper is swinging pendulum-wise between the poles of an electromagnet, (*b*) when the current flowing through the coils of an electromagnet is suddenly interrupted, (*c*) when a starting-resistance is used for a motor.

192. The coil of a moving-coil galvanometer is short-circuited when the galvanometer is swinging. Describe and explain what happens.

193. Describe experiments illustrating—

(1) The essential conditions for the production of an induced E.M.F.

(2) The factors upon which the direction of the induced E.M.F. depends.

(3) The factors upon which the magnitude of the induced E.M.F. depends.

194. Explain how the laws of electromagnetic induction are illustrated in the case of the transformer.

195. Give a short account of the general arrangement of a D.C. motor. What is meant by the back E.M.F. of a motor?

196. Describe the essentials of a dynamo and explain its action. Explain what factors determine the E.M.F. and efficiency, why the armature coils are wound on a soft-iron core, and how the commutator works.

197. Explain the induction coil in detail, showing clearly how it can produce very high voltages. On what factors does the voltage depend?

198. A dynamo is charging accumulators, when the driving belt breaks. It continues to run, being driven as a motor by current from accumulators. Will it be revolving in the same direction as before (*a*) if it is series wound, (*b*) if it is shunt wound?

199. What is meant by self-inductance? Describe experiments to illustrate it. In what units is it measured, and how are they defined?

Numerical Questions

Tangent Galvanometers and relations between steady currents and magnetic fields

200. Find the magnetic field at the centre of a circular coil of radius 5 cm. carrying a current of 4 amperes.

201. A tangent galvanometer with 50 turns of a resistance of 4 ohms is deflected through 60° by an applied voltage of 0·3 volt. What is the radius of the coil if H = 0·18 oersted? What is the reduction factor of the galvanometer?

202. How many turns of wire of radius 10 cm. would you wind on a tangent galvanometer which was required to give a deflection as near as possible to 45° for 1 milliamp. where H was 0·18 oersted.

203. A tangent galvanometer contains two concentric coils, one of 10 turns of radius 10 cm. and the other of 30 turns of radius 15 cm. Find its reduction factor if H = 0·18 oersted.

204. Deduce the formula for the *sine galvanometer* in which the coil is rotated until the needle lies in its plane. What deflection on a sine galvanometer corresponds to a deflection of 30° on a tangent galvanometer of the same dimensions?

205. Find the force on a straight wire a metre long carrying 20 amperes at right angles to a magnetic field of 40 oersted.

206. A wire 2 metres long hangs vertically with a current of 100 amperes passing down it. If the declination is 15° W., what force acts on the wire and in what direction? (H = 0·18 oersted.)

207. The moving coil of a galvanometer has 100 square turns of side 2 cm. The field-strength is 2000 oersted.

Find the couple for a current of 1 milliamp.

208. A point-pole has strength 500 webers. Find (*a*) the total flux leaving it; (*b*) the magnetic field 5 cm. away; (*c*) the flux across 5 sq. cm. of the surface of a sphere of 10 cm. radius surrounding it.

209. The flux in the air-gap of an electromagnet is 50,000

lines, and the area of the surfaces of the pole-pieces is 10 sq. cm. Find the pole-strength and average magnetic field in the gap.

210. A wire carries a current of 10 amperes at right angles to the lines of force between the poles of strength 100 webers and of area 20 sq. cm. Find the force on 2 cm. length of the wire, assuming the field to be uniform.

211. A vertical wire carries a current of 5 amperes in the earth's magnetic field. $H = 0{\cdot}18$ oersted. In what direction and at what distance will a neutral point be found?

212. Two long straight parallel wires carry currents of 10 amps. 5 cm. apart. What is the force between them per cm. length?

Induced Currents

213. What is the average E.M.F. and current when a wire, part of a circuit of resistance 10 ohms, is drawn in $\frac{1}{20}$ of a second across the air-gap of an electromagnet of pole strength 5000 webers?

214. A wire slides on frictionless rails. The wire, the rails, and a magnetic field of 1000 oersted are mutually at right angles. There is a P.D. of 10 volts between the rails. The resistance of the wire and its contacts is negligible. The wire is 20 cm. long. What final speed will it attain?

215. In the apparatus of Question 214, the rails are connected by a wire of resistance 3 ohms, and no P.D. is maintained between the rails. What force will be required to give the wire a velocity of 100 cm./sec.?

216. A disc of radius 20 cm. revolves at 100 revs. per sec. in a field of 1000 oersted. The surface is in contact with a mercury cup. What is the P.D. between the axle and the mercury?

217. A coil has an inductance of 2 henries. What average E.M.F. is induced when the current in the coil changes from zero to 50 amperes in $\frac{1}{100}$ second?

218. An inductance of 20 henries carries 100 milliamps. Its resistance is 80 ohms. A non-inductive resistance of 120 ohms is in parallel with the coil. The external circuit is suddenly broken. What charge circulates?

219. A coil of 200 turns, area 10 sq. cm., resistance 5 ohms, revolves 1000 times a second about an axis perpendicular to a field of 2000 oersted. A commutator-system keeps the current all in the same direction. What is its average value?

220. A flux of 20,000 lines threads a coil of 100 turns. $\frac{1}{50}$ of

a second later there are 30,000 lines. What is the average of E.M.F. in volts?

221. A telegraph wire running magnetic east and west where H = 0·18 oersted is blown down. Successive posts are 60 metres apart. The wire falls freely from a height of 5 metres. Taking g as 1000 cm./sec.^{2}, find the average voltage induced between two posts during the fall.

222. An earth inductor having 500 turns of area 200 sq. cm. and resistance 10 ohms makes 100 revs. per sec. in the earth's field of 0·5 oersted. A commutator is in the circuit. What average current flows?

223. If the vertical component of the earth's field is 0·45 oersted, find the E.M.F. induced in the axle, 150 cm. long, of a railway carriage travelling at 90 kilometres per hour.

Dynamos and Motors

224. Each conductor in an armature of a motor is 50 cm. long and carries 20 amperes. The radial magnetic field is 20,000 oersted, and 500 conductors are under the poles at any instant. The radius of the armature is 40 cm. Find the couple acting.

225. The armature resistance of a shunt-wound dynamo is 0·3 ohm, and the field resistance 50 ohms. It supplies 30 amps. to an external circuit at a p.d. of 200 volts. Find the lost volts in the armature, and the e.m.f. of the dynamo.

226 The external current in Question 225 is increased to 50 amps., and some resistance is cut out of the field to keep the field current the same. What is now the supply P.D.? How much resistance was cut out of the field circuit?

227. A shunt motor takes 20 amps. at 200 volts. Its armature resistance is 0·2 ohm and its field-resistance 40 ohms. Find the back E.M.F.

228. The load is doubled in the motor of Question 227. Find the new back E.M.F. and the ratio of the new revs. to the former revs.

229. The field-resistance in Question 227 is increased to 50 ohms, the load being unchanged. Find the ratio of the new speed to the old speed.

230. The load in Question 227 is reduced to two-thirds of its former value. Find the ratio of the new speed to the old speed.

231. What starting resistance is needed by a series motor of resistance ½ ohm running off 100-volt mains, if the current must

not exceed 20 amps.? What is the back E.M.F. when it is running with this current?

232. A series motor of resistance 1 ohm takes 20 amps. when generating 2 kilowatts. Find the back E.M.F. and supply P.D. Efficiency 40%.

233. A series motor of resistance 2 ohms runs at 1000 r.p.m. when loaded so as to take 20 amps. from 100-volt mains. If it is unloaded till the current is 10 amps., what speed will it settle down at?

234. A separately excited dynamo of E.M.F. 110 volts and armature resistance 0·1 ohm is in parallel with a bank of accumulators of E.M.F. 100 volts and resistance 0·05 ohm. The combination supplies current to an external circuit. Find the current supplied by the dynamo, and the P.D. of supply, when the combination supplies (*a*) 50 amps., (*b*) 100 amps., (*c*) 200 amps.

PART II

Die Mathematiker sind eine Art Franzosen: redet man zu ihnen, so übersetzen sie es in ihre Sprache, und dann ist es alsbald ganz etwas anderes.

GOETHE: " Aphorismen über Naturwissenschaft."

The mathematician is a kind of Frenchman. You tell him something; he translates it into his own language; and at once it is something else.

CHAPTER I

THE FOUNDATIONS OF ELECTROSTATIC THEORY

Basic Assumptions—The Inverse Square Law—Dielectrics—Intensity, Potential and Displacement—Potential due to Point-charge—Potential Uniform on Electrostatic Conductor—Potential of Charged Sphere in Free Space—Capacity—Parallel Plate Condenser—Spherical Shell Condenser—Cylindrical Shell Condenser—Use of Gauss's Theorem—Energy of a Charged Conductor or Condenser—Energy of a Parallel Plate Condenser—Energy stored in the Dielectric—Electrical Images.

Basic Assumptions

ELECTRIC charge is whatever gives amber rubbed with fur the power of attracting light objects and repelling other pieces of rubbed amber. Electric charge also appears on glass rubbed with silk.

Charge on amber rubbed with fur is found to attract charge on glass rubbed with silk. Charge on amber rubbed with fur is defined as negative charge, and charge on glass rubbed with silk is defined as positive charge. Like charges thus repel, and unlike charges attract.

A conductor is defined as a substance which allows the passage of electricity; an insulator as a substance which resists the passage of electricity. It is clear from these definitions that no rigid line can be drawn between conductors and insulators. In electrostatics it is assumed that the distribution of potential and charge, on a conductor in an electric field, reaches its final state in an indefinitely short time after the field is applied.

According to the modern theory, negative charge in solids

is carried by electrons, which are themselves minute electrical charges. Positive charges are carried by atomic nuclei, which do not move about in solids. On the modern theory, therefore, negative charges alone move on solid conductors. A negative charge on a solid is caused by an excess of electrons; a positive charge by a shortage of them. The fundamental relations between charge, potential, intensity, capacity, and energy are the same whichever theory is used, provided that we are considering large-scale phenomena rather than individual atoms. Since the fundamental relations were originally founded on the assumptions that electric charge is continuous rather than divided into small particles, and that both positive and negative charges move freely in conductors, we shall use these assumptions here.

The Inverse Square Law

Coulomb's Torsion Balance experiment (see p. 14) gave a strong presumption, by direct evidence, that two charges exerted on each other a force varying inversely as the square of the distance between them. Cavendish's experiment (see p. 14) confirmed this with great exactness, for it is possible to show not only that the results of Cavendish's experiment would follow if the Inverse Square Law were true (this follows from Gauss's Theorem, given in the Appendix), but also that if Cavendish's experiment were accepted the Inverse Square Law alone could account for its results (Cavendish's Proof, given in the Appendix). Repeating Cavendish's experiment with improved apparatus, Maxwell showed that it followed from his own result that the force between two charges a distance r apart must vary as

$$\frac{1}{r^{\left(2 \pm \frac{1}{216,000}\right)}}$$

We are therefore justified in accepting the Inverse Square Law.

"Unit charge on the electrostatic system (sometimes called one statcoulomb) is therefore the charge which exerts a force of 1 dyne on an equal charge 1 cm. away *in vacuo*, both charges being assumed to be concentrated at points."

Dielectrics

The force between two charges in a material medium is less than between the same charges the same distance apart *in vacuo*. It is assumed that one charge acts on another, whether *in vacuo* or in a material medium, through what is called the "Dielectric," which may be regarded as whatever occupies the space between the charges.

The ratio of the force *in vacuo* to the force in the medium is called the "Dielectric Constant." Thus the dielectric constant *in vacuo* is 1.

The force between charges Q_1 and Q_2, situated r cm. apart in a medium of dielectric constant k, is then

$$\frac{Q_1 Q_2}{kr^2} \text{ dynes.}$$

Intensity, Potential and Displacement

The Intensity at a point is numerically equal to, and in the direction of, the force in dynes on a unit positive charge at the point. It will be called E.

The Potential at a point is numerically equal to the work done in ergs in bringing unit positive charge from infinity to the point, against the Intensity. It will be called V.

The Displacement, D, is the number of tubes of induction per square centimetre, a "tube" containing all the lines of force leaving unit charge. As 4π lines leave unit charge, we have $D = \frac{kE}{4\pi}$. (See p. 32.)

Let us consider the relation between E and V.

The work done in bringing unit charge from any point to a neighbouring point dr away is $-\,Edr$.

The negative sign is necessary, since, in order to do positive work against the intensity, E must act in a direction opposite to that in which r is increasing.

We have therefore, if we are considering a point r cm. from an assumed origin,

$$V = \int_{\infty}^{r} -\,Edr.$$

Differentiating this equation to find E, we get

$$E = -\frac{dV}{dr},$$

which is the mathematical form of the statement that "The Intensity is the negative gradient of the Potential."

There is a subtle difficulty about this equation. It assumes that we are taking the change of V in the particular direction in which it is changing most quickly. If we wanted a component of E in another direction we should still have to differentiate. But we should then want a partial differential coefficient only, and should therefore write

$$E = -\frac{\partial V}{\partial r}$$

with some subscript specifying the particular direction we were choosing. In this elementary work we shall say $-\frac{dV}{dr}$ when we mean that we are taking the direction in which the potential is falling most quickly.

It may be helpful to visualize the electric field round a point charge as represented by a big indiarubber sheet, fixed at its edges to a horizontal plane, with its middle pushed up from underneath by a vertical pin which does not pierce it. The slope of the stretched sheet is greatest near the point, and falls off as we approach the edge. On this sheet the tangent of the greatest angle of slope at any point might represent the intensity, the direction in which the slope is greatest the direction of the intensity, and the vertical height above the horizontal plane the potential. It is then clear that the intensity is the negative gradient of the potential.

Potential due to Point-charge

The intensity at a distance of r cm. from a point-charge is $\frac{Q}{kr^2}$.

$$\therefore \quad V = \int_{\infty}^{r} -\frac{Q}{kr^2}dr$$

$$= \frac{Q}{kr}.$$

The constant of integration is zero, if, as is usual, the potential at infinity is taken as zero.

Potential Uniform on Electrostatic Conductor

Since the charge is not moving, E is zero everywhere.

$$\therefore \quad \frac{dV}{dr} = 0.$$

Integrating,

$$V = \text{constant}.$$

Potential of Charged Sphere in Free Space

The distribution of the lines of force shows that the intensity outside the sphere is everywhere the same as if the whole charge had been concentrated at the centre. If the sphere has radius r and charge Q, in medium of dielectric constant k, we have for V at the surface (which is the same as V throughout the sphere)

$$V = \int_{\infty}^{r} -\frac{Q}{kr^2}dr$$

$$= \frac{Q}{kr}.$$

Capacity

Capacity is defined as the ratio $\frac{\text{Charge}}{\text{Potential}}$.

Thus the capacity of a conducting sphere in free space is

$$\frac{Q}{Q/kr}, \text{ or } kr.$$

In vacuo, therefore, the capacity of a sphere is numerically equal to its radius.

Parallel Plate Condenser

If we neglect end-corrections, and take σ as surface density of charge and A as the area of plates, distance d cm. apart, in medium of dielectric constant k, the intensity between the plates is $\frac{4\pi\sigma}{k}$. (See p. 31.)

$$\therefore \quad V = \int_{d}^{0} -\frac{4\pi\sigma}{k}dr = \frac{4\pi\sigma d}{k}.$$

Since the total charge is σA

$$C = \frac{\sigma A}{\frac{4\pi\sigma d}{k}} = \frac{kA}{4\pi d}. \qquad \text{(compare proof on p. 35)}$$

If the plates are separated partly by a slab of dielectric constant k, t cm. thick, and are otherwise *in vacuo*—

$$\begin{aligned} V &= \int_t^0 -\frac{4\pi\sigma}{k}dr + \int_{(d-t)}^0 -4\pi\sigma d \\ &= \frac{4\pi\sigma t}{k} + 4\pi\sigma(d-t) \\ &= 4\pi\sigma\left\{d - \left(1 - \frac{1}{k}\right)t\right\}. \end{aligned}$$

So the capacity is $$\frac{A}{4\pi\left\{d - \left(1 - \frac{1}{k}\right)t\right\}}.$$

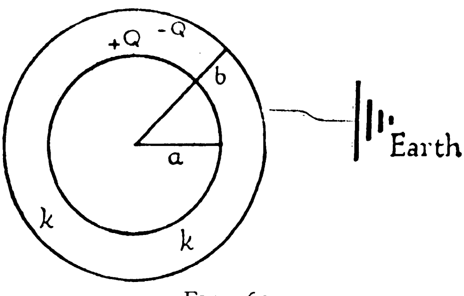


FIG. 162.

Spherical Shell Condenser

Let the outer shell, of radius b, be earthed, and the inner one, of radius a, have a charge Q. A medium of dielectric constant k separates the shells (see Fig. 162). By symmetry, the lines of force, and therefore the intensity, between the spheres are the same as those round a point charge Q at the centre, except that they are cut off everywhere outside and inside two spherical shells of radii a and b. The potential of the charged sphere is thus

$$V = \int_b^a -\frac{Q}{kr^2}dr$$

$$= \frac{Q}{k}\left(\frac{1}{a} - \frac{1}{b}\right).$$

So the capacity
$$= \frac{Q}{\frac{Q}{k}\left(\frac{1}{a} - \frac{1}{b}\right)}$$

$$= \frac{k}{\frac{1}{a} - \frac{1}{b}}$$

$$= \frac{kab}{b - a}.$$

Cylindrical Shell Condenser

Consider now the case of a cylindrical condenser, of length l and internal and external radii a and b (see Fig. 163). Assume

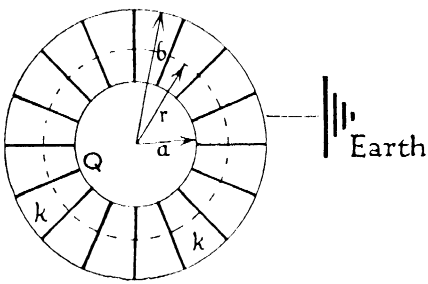


FIG. 163.

also that it has guard-ring condensers at both ends to cut out end-corrections, as they were cut out for the parallel plate condenser. In this case we will assume a vacuum between the plates.

The charge Q on the inner cylinder induces an equal and opposite charge on the outer cylinder.

Since Q tubes of force leave the inner cylinder and arrive on the outer one, $4\pi Q$ lines of force cross any cylindrical surface (of radius r) between the inner and outer ones.

Hence the number of lines of force crossing per sq. cm. of surface of radius $r = \dfrac{4\pi Q}{\text{Area of cylindrical surface of radius } r}$.

But the number of lines of force per sq. cm. is the intensity *in vacuo*. Hence the intensity E at a point r from the centre is given by

$$E = \frac{4\pi Q}{2\pi r l} = \frac{2Q}{rl}.$$

[This is an important result in itself.]

Hence potential of inner cylinder

$$= \int_b^a - \frac{2Q}{rl} dr = - \frac{2Q}{l} (\log a - \log b) = \frac{2Q}{l} \log \frac{b}{a}.$$

$$\text{Hence } C = \frac{Q}{\frac{2Q}{l} \log \frac{b}{a}} = \frac{l}{2 \log \frac{b}{a}}.$$

If there were a medium of dielectric constant k between the plates, the capacity would clearly become

$$\frac{kl}{2 \log \frac{b}{a}}.$$

Use of Gauss's Theorem

All capacity problems are perhaps better attacked by the direct use of Gauss's Theorem (see p. 485).

All depend upon the determination of the potential V, and this in turn depends on the intensity E.

We have up to now deduced the intensity entirely from the consideration of lines of force, which is really rather an unwieldy way of doing it, though easy to understand.

Gauss's Theorem, however, gives us the intensity directly.

This Theorem states that the total normal induction across any surface is $4\pi \times$ the charge inside.

The normal induction is the sum of the products of all the normal components of the intensity E across the surface considered, multiplied by the area elements these components cross, and by k. If we consider only surfaces such that the intensity is everywhere normal to them, the normal induction will simply be the product of the intensity, the area of the surface, and k.

Let us now find the intensity E at a distance r from a sphere of radius a ($<r$) in a medium of dielectric constant k, carrying

charge Q. Let our Gauss surface be a sphere of radius r concentric with the charged sphere.

Then by Gauss's Theorem—

$$kE \times \text{area of gauss surface} = 4\pi Q.$$

$$\therefore \quad kE \times 4\pi r^2 = 4\pi Q$$

$$E = \frac{Q}{kr^2} \text{ as before.}$$

This result also applies to the spherical shell condenser, provided that $a < r < b$.

For the cylindrical shell condenser, let us take as our gauss surface a cylinder of length l and radius r, so that $a < r < b$.

Then as before

$$kE \times 2\pi rl = 4\pi Q$$

$$E = \frac{2Q}{krl} \text{ as before.}$$

In the case of the parallel plate condenser, let us take our gauss surface as the surface of a cube of side a, so that $a < d$, some of it inside the earthed plate and some outside, its outside face being parallel to the earthed plate. Then if we are not near enough to the edge for edge-effects to distort the field,

$$kE \times a^2 = 4\pi\sigma a^2$$

$$E = \frac{4\pi\sigma}{k} \text{ as before.}$$

Energy of a Charged Conductor or Condenser

The energy of the charged conductor is equal to the work done in charging it.

The work done in bringing an element of charge dQ to the conductor when its potential is V is, by the definition of V, VdQ ergs.

Hence, if W be the energy,

$$W = \int_0^Q V dQ \text{ ergs.}$$

This cannot be integrated directly, since V is not constant, but is proportional to Q; for $V = \frac{Q}{C}$.

$$\therefore \quad W = \int^{Q} \frac{Q}{C} dQ$$

$$= \frac{1}{2}\frac{Q^2}{C}.$$

Since $Q = CV$, and $C = \frac{Q}{V}$, alternative expressions for the energy are—

$$W = \tfrac{1}{2}CV^2$$

and

$$W = \tfrac{1}{2}QV.$$

Energy of a Parallel Plate Condenser

We have already seen on p. 45 that the mechanical force on the surface of a conductor is $\frac{2\pi\sigma^2}{k}$ per unit area.

The work done in separating the plates of a parallel plate condenser is then

$$\frac{2\pi\sigma^2 Ad}{k}$$

But the energy

$$= \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2} \cdot A^2\sigma^2 \cdot \frac{4\pi d}{kA}$$

$$= \frac{2\pi\sigma^2 Ad}{k}.$$

So the energy of a parallel plate condenser is the same as the mechanical work done in separating the plates. This may be regarded as confirming our formulæ.

Energy stored in the Dielectric

The mechanical force per sq. cm. of surface on a charged conductor is, as we have seen (p. 44)—

$$\frac{2\pi\sigma^2}{k}.$$

The intensity E at the surface is given by

$$E = \frac{4\pi\sigma}{k}$$

so

$$\sigma = \frac{kE}{4\pi}.$$

Substituting in the expression $\frac{2\pi\sigma^2}{k}$ we get mechanical force per unit area—

$$= \frac{2\pi}{k} \frac{k^2E^2}{16\pi^2}$$

$$= \frac{kE^2}{8\pi}.$$

The work done in moving a charged surface of area 1 sq. cm. through a distance of 1 cm. along its normal is thus

$$\frac{kE^2}{8\pi} \text{ ergs.}$$

This work has been done in sweeping out 1 c.c. of dielectric, and setting up an electric field in it. No other change has been effected. The work has then been used in putting the dielectric in a state of strain.

Energy to the amount of $\frac{kE^2}{8\pi}$ has thus been stored in 1 c.c. of dielectric.

In general the energy per unit volume stored in an electric field is thus

$$\frac{kE^2}{8\pi} \text{ ergs.}$$

It is interesting to note that since the Electric Displacement D is given by $D = \frac{kE}{4\pi}$, we may say—

$$\text{Energy per unit Volume} = \frac{DE}{2}.$$

Electrical Images

We have seen in Part I, p. 28, how we may consider generally the effect of the introduction of an uncharged body in an electric field on both the intensity and the potential in the neighbourhood.

It is worth while to give a little more consideration to the kind of problem raised in this case. We have already quite a number of facts, or theorems, either explicitly stated or implied, to help us.

The explicitly stated facts or theorems are these:

(*a*) The intensity in any direction is the negative potential gradient in that direction.

(*b*) The potential is constant, and the intensity zero, at every point in an electrostatic field occupied by conducting material.

(*c*) The direction of the intensity is normal to every equipotential surface, and every surface of conducting material. In other words, lines of force leave a conducting surface, and cross an equipotential surface, normally.

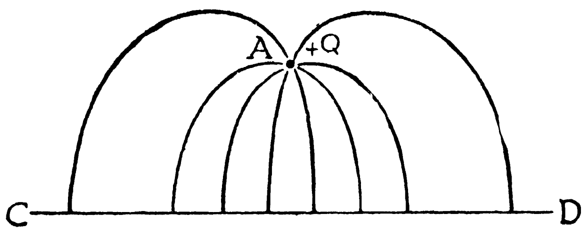


FIG. 164.

(*d*) Lines of force can only begin on positive charges, and end on negative charges.

Implied facts are—

(*e*) The potential cannot have a maximum or minimum at any point not occupied by a charge.

(*f*) Wherever the potential has a maximum there is a positive charge, and where it has a minimum there is a negative charge.

We can in this way get quite a good qualitative impression of the behaviour of the potential and intensity for any given arrangement of charges and conductors. For instance, we can see that, if we have a point charge near an infinite conducting plane, the lines of force will go more or less as in Fig. 164, so that the space rate of change of intensity will diminish as the plane is approached; and the surface-density of induced negative charge will be greatest at the foot of the perpendicular from the point to the plane.

The calculation of the actual intensity and potential at any given point will, however, be a very arduous business, to be left to the mathematicians for solution, unless we can find an easy way.

There is, however, a quite easy way, known as the method of electrical images.

Consider in the above case that the infinite conducting plane is replaced by a sheet of zero thickness at its surface, and another charge of $-Q$ at B, so that if AN is perpendicular to the plane, ANB is a straight line in which AN = NB. Every point of the surface CD is now at zero potential, as

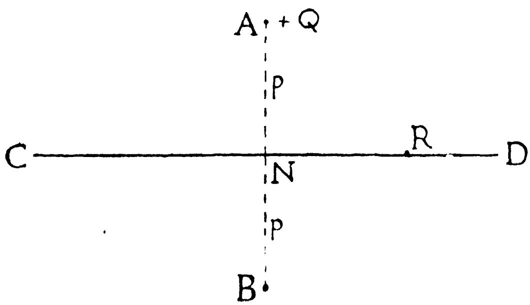


FIG. 165.

before, and CD is again an equipotential surface. The distribution of lines of force, and therefore of intensity and potential, near CD is now exactly the same as in the real case, but now they are easy to calculate, since they are due solely to two equal and opposite point charges. See Fig. 165.

The complete proof of this latter statement, though not difficult, is rather long and of more interest to the pure mathematician than the physicist, and it may be found in any standard work on the subject.

Let us, however, calculate the potential at any point, the surface density of induced charge at any point on the surface CD, and the intensity at any point on the surface.

Let N be the origin of co-ordinates, with AB along the y-axis and CD along the x-axis.

Let A be the point $(0, p)$, and B the point $(0, -p)$.

Consider any point P whose co-ordinates are (x, y).

Then the potential at $P = \frac{Q}{PA} - \frac{Q}{PB}$.

$$= \frac{Q}{\sqrt{x^2 + (p - y)^2}} - \frac{Q}{\sqrt{x^2 + (p + y)^2}}.$$

The intensity at a point R, whose co-ordinates are $(r, 0)$,

$$= 2\frac{Q}{AR^2} \cdot \frac{AN}{AR}$$

$$= \frac{2pQ}{(p^2 + r^2)^{\frac{3}{2}}}.$$

If σ is the surface-density of charge at R,

$$E = 4\pi\sigma$$

$$\sigma = \frac{pQ}{2\pi(p^2 + r^2)^{\frac{3}{2}}}.$$

These three results show the remarkable simplicity of the method.

CHAPTER II

ELECTROMAGNETISM. MAGNETIC SHELL METHOD

Magnetic Shells—Solid Angles—Potential due to a Short Bar-magnet —Potential due to a Magnetic Shell—Ampère's Theorem—Definition of Unit Current—Intensity on the Axis of a Circular Coil—Intensity on the Axis of a Solenoid—Intensity on the Axis of a Long Solenoid—Work done in taking Unit Pole round a Current—Line Integral of a Magnetic Field—Intensity near a Long Straight Current—Couple acting on a Coil in a Magnetic Field.

Magnetic Shells

THE conception of the Magnetic Shell is of the very greatest importance in electromagnetic theory, since the definition of the electromagnetic unit of current is based on it, and many problems, concerned with the magnetic fields due to currents, are solved by it.

Imagine a thin sheet of magnetic material magnetized everywhere in a direction normal to its surfaces. North pole may be imagined to be uniformly spread over one surface, and south pole over the other. The magnetic moment per unit area of such a sheet would be the product of the pole-strength per unit area and the thickness.

Let us now consider that the sheet is indefinitely thin, and let σ be its magnetic moment per unit area. We may leave both the pole-strength per unit area and the thickness indeterminate, while fixing their product as σ.

This is the conception of the Magnetic Shell. All we can say definitely about its thickness is that in any problem the thickness is to be regarded as so small as to be quite negligible for that particular problem, though it can never be zero, since σ must have a finite value.

σ, the magnetic moment per unit area, is known as the "strength" of the shell.

Magnetic shells can be shown to have a most remarkable property. The magnetic intensity at any external point due

to a magnetic shell can be shown to depend only on the boundary of the shell and the strength, and to be completely independent of the shape. The exact theorem which is about to be proved about magnetic shells may be stated as follows :

"*If a magnetic shell of strength σ subtends a solid angle ω at a point P, the magnetic potential at P due to the shell is $\sigma\omega$.*"

Two conceptions in this theorem must be explained in detail —*solid angle*, and *magnetic potential.*

Solid Angles

Consider the circle, centre A, in Fig. 166.

We know that the angle subtended at A by the arc or chord BC is, in radian measure, equal numerically to the actual length of the arc DE cut off by AB and AC on the small dotted circle of unit radius round A. Thus if $\angle$BAC is a right angle the length of this arc is $\frac{\pi}{2}$, which is the magnitude of $\angle$BAC in radian measure.

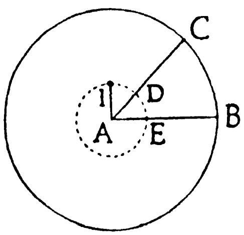


FIG. 166.

The sum of all the plane angles round a point is thus 2π, the circumference of a circle of unit radius round the point.

The same kind of conception is used for " solid " angles. If we wish to find the solid angle subtended by a triangle BCD at a point A not in the plane of BCD, we imagine a unit sphere described round A. The total area of its surface is 4π units. Let us now imagine the area cut off on the surface of this unit sphere by straight lines joining A to every point on each side of the triangle; that is, to every point on the triangle's perimeter. [We may also, if we like, consider it as the area cut off on the surface of the sphere by the three planes ABC, ACD, ADB.]

This area is numerically equal to the solid angle subtended at A by the triangle BCD.

The conception of the solid angle is of the highest importance in Physics. For example, the proportion of the total radiation emitted by any point source which falls on a

body in its neighbourhood obviously depends on the solid angle subtended by the body at the source. If the earth subtends a solid angle ω at the centre of the sun (for all the sun's light radiates in straight lines which would have passed through its centre if they had been produced backwards), then the earth receives a fraction $\frac{\omega}{4\pi}$ of the sun's total radiation.

Let us now consider the actual value of the solid angle subtended by a small element of area δA at a distance r from a point P.

For simplicity, let δA have its plane perpendicular to the straight line of length r joining δA to P. Let it subtend a solid angle $\delta\omega$ at P (Fig. 167).

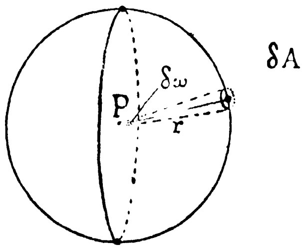


FIG. 167.

Now the solid angle subtended by the whole sphere round P is 4π, since 4π is the area of the unit sphere round P, just as 2π was the length of the circumference of the unit circle round A in Fig. 166.

Then we have

$$\frac{\delta\omega}{4\pi} = \frac{\delta A}{\text{area of sphere radius } r}$$

$$= \frac{\delta A}{4\pi r^2}.$$

Hence $$\delta\omega = \frac{\delta A}{r^2}.$$

Now let δA, still remaining a small element of area, be tilted as in Fig. 168 so as to make an angle θ with the surface-plane of the sphere in the neighbourhood. The solid angle subtended by δA at P will now be that subtended by its projection $\delta A'$. Now the breadth of $\delta A'$ is equal to the breadth of δA, but the height of $\delta A'$ is $\cos\theta \times$ the height of δA. So we have

$$\delta A' = \delta A \cos\theta.$$

Now $$\delta\omega = \frac{\delta A'}{r^2}$$

$$= \frac{\delta A \cos\theta}{r^2}.$$

[It must be realized that this proof only holds when δA is an *element* of area. If it had a finite value we should have to consider the curvature.]

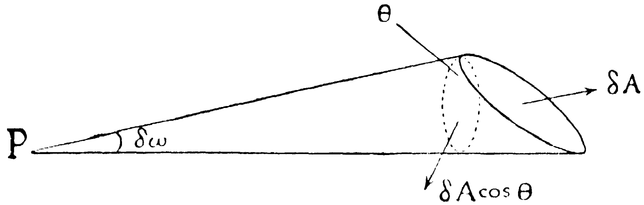


FIG. 168.

Potential due to a Short Bar-magnet

Let NS in Fig. 169 be a short bar-magnet of pole-strength m and length $2l$. Let AB be its perpendicular bisector. Let O be its mid-point. It is required to find the potential due to it at an external point P. Let OP $= r$, and $\angle$SOP $= \theta$.

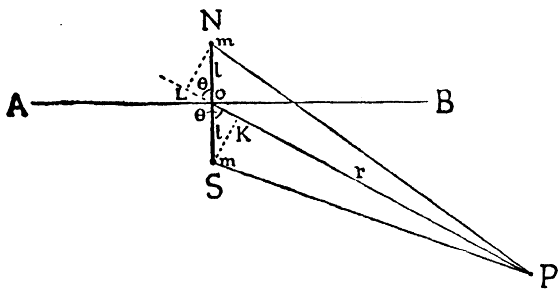


FIG. 169.

Now it follows from the Inverse Square Law that the electric potential at a distance r from a charge Q is $\frac{Q}{r}$. Magnetic poles also obey the Inverse Square Law. So the potential at

$$P = \frac{m}{PS} - \frac{m}{PN}.$$

Since NS is a short magnet, l is very small compared with r. Let SK and NL be perpendiculars dropped from N and S on PO.

If we neglect small quantities of the second order, PS = PK, and PN = PL.

Also OL = OK = $l \cos \theta$.

So the potential at P

$$V = \frac{m}{r - l \cos \theta} - \frac{m}{r + l \cos \theta}$$

$$V = \frac{2ml \cos \theta}{r^2 - l^2 \cos^2 \theta}$$

$$V = \frac{M \cos \theta}{r^2},$$

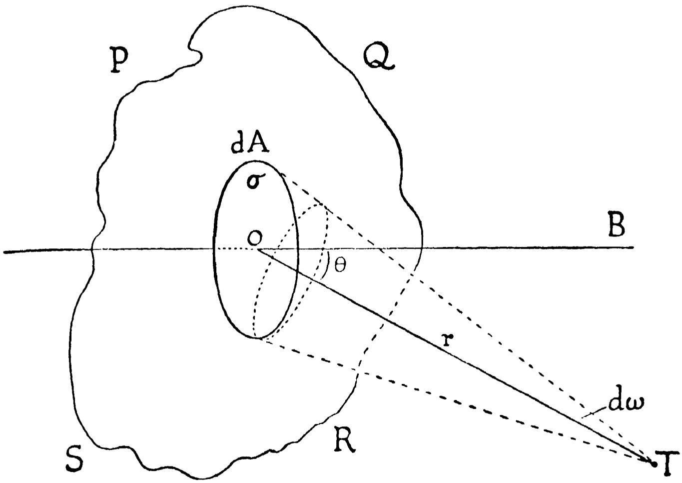


FIG. 170.

if we put M, the magnetic moment, for $2ml$, and neglect $l^2 \cos^2 \theta$ by comparison with r^2.

As we shall soon be considering bits of an "indefinitely thin magnetic shell" these approximations do not introduce any inaccuracy.

We are now in a position to prove the magnetic shell theorem.

Potential due to a Magnetic Shell

Let PQRS (Fig. 170) be a magnetic shell of strength σ and area A. Let it subtend a solid angle ω at a point T. Let dA,

an element of area of A, subtend a solid angle $d\omega$ at T. Let O be a point in dA, and let OB, the normal to the shell at O, make an angle θ with OT.

Then from the discussion of the potential due to a short bar-magnet we see that the element of potential dV at T due to the element of shell dA, having magnetic moment σdA, is given by

$$dV = \frac{\sigma dA \cos\theta}{r^2}.$$

But from the discussion of solid angles $\frac{dA \cos\theta}{r^2} = d\omega$.

Hence
$$dV = \sigma d\omega,$$
$$V = \sigma \int d\omega$$

where the integration is taken over the shell.

$$\therefore \quad V = \sigma\omega,$$

which proves the theorem.

Ampère's Theorem

We can now consider Ampère's Theorem. It states that:

" *A linear conductor carrying a current is equivalent to a magnetic shell whose boundary coincides with the linear conductor.*"

There are still some hard words. " Linear " means " of such small cross-section that the cross-section may be neglected in all calculation." " Equivalent " means " producing the same magnetic potential (or intensity) at all external points."

The general justification of this theorem is that the system of measurement on which it is based is found to agree with experience.

The particular justification for taking the magnetic effect, rather than (say) the heating effect, to define the magnitude of a current, is in Rowland's experiment, which shows that the magnetic intensity due to a moving electrostatic charge is proportional to the rate of passage of charge. This experiment is described in the early part of Chapter VIII, Part I (p. 108).

The particular justification for Ampère's Theorem is that direct experiment shows that a circuit of small area is

equivalent at external points to a small magnet with its magnetic axis normal to the plane of the circuit. The "magnetic moment" of the circuit is found to be proportional to the area of the circuit and the rate of flow of charge round it.

Definition of Unit Current

We may now use Ampère's Theorem to give us the absolute electromagnetic unit of current.

It may be defined as follows:

"*The Absolute Electromagnetic Unit of current is that current which is equivalent to a magnetic shell of unit strength having the same boundary.*"

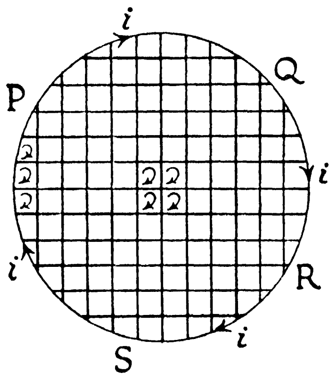


FIG 171.

The odd thing about this definition is that it allows us to give our magnetic shell any shape we like, provided that its boundary is the current.

This may perhaps be better understood by a consideration of Fig. 171. Let PQRS be a circuit carrying a clockwise current of i absolute electromagnetic units. Let a complete network of wires fill the circuit as in the figure, and let the current i flow round each of them clockwise.

Each bit of the outer circuit still has its current i, so the outer circuit is unchanged. Each complete circuit in the network has its circular current. As these elements may be made as small as we like, we may more easily imagine each replaced by an element of magnetic shell with the same boundaries.

Now if one looks at any single wire in the network, one sees that two currents flow in it, equal and opposite. So the total current is zero. So putting in the network has made no difference to the situation, except that it has given our imagina-

tions a legitimate aid. We can see from this directly that if Ampère's Theorem ever works, then only the boundary of the current matters. The shape we choose for the shell does not affect the result. This consideration is not a formal proof; only an aid to the imagination.

Intensity on the Axis of a Circular Coil

Let P be a point on the axis of a circular coil AB (Fig. 172) AB is seen from a point in its own plane in this figure. Let R be the radius of the coil AB, x the distance of its circumference from P, and y the distance of its centre from P. Let C be a point on the axis beyond P, so that PC $= x$, and let D be the centre of the coil AB. Let AB carry a current i.

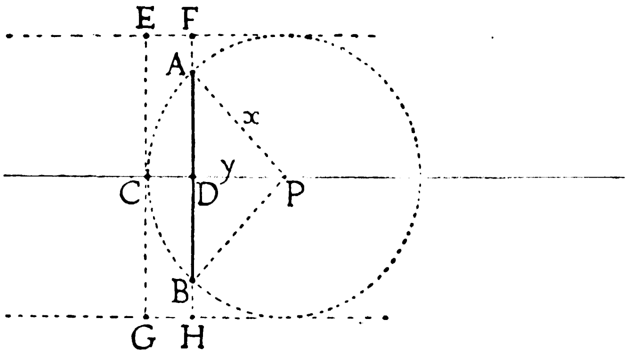


FIG. 172.

To find the Magnetic Potential at P due to the coil, we replace the coil by a magnetic shell, of strength $\sigma = i$, having the coil AB as a boundary, but having any shape we like.

Obviously the best shape to choose is one which puts all the shell at the same distance from P, and thus simplifies as much as possible the calculation of the magnetic potential. Such a shape is the part of a sphere of radius x and centre P cut off by the coil AB.

Now the area of a segment of a sphere cut off between two parallel planes is equal to the area cut off on the sphere's circumscribing cylinder by the same two planes. [If there is

any doubt about this it should be proved. The proof is simple.]

Let EF be the top, and GH the bottom, in the figure, of the circumscribing cylinder cut off by the planes through C and D perpendicular to CP.

The area of the cylinder EFGH, and therefore of the segment ACB, is given by

$$\begin{aligned} \text{area} &= 2\pi x \times \text{CD} \\ &= 2\pi x(x - y). \end{aligned}$$

$\therefore$ Solid angle subtended at P by the coil AB is given by

$$\begin{aligned} \omega &= \frac{2\pi x(x - y)}{x^2} \\ &= 2\pi\left(1 - \frac{y}{x}\right). \end{aligned}$$

So the magnetic potential

$$\begin{aligned} \text{V} &= 2\pi i\left(1 - \frac{y}{x}\right) \\ &= 2\pi i\left(1 - \frac{y}{\sqrt{y^2 + \text{R}^2}}\right). \end{aligned}$$

Now by symmetry the intensity H is along the axis. So we have

$$\begin{aligned} \text{H} &= -\frac{d\text{V}}{dy} \\ &= -\frac{d}{dy}\left\{2\pi i\left(1 - \frac{y}{\sqrt{y^2 + \text{R}^2}}\right)\right\} \\ &= 2\pi i\left\{\frac{\sqrt{y^2 + \text{R}^2} \times (1) - y \cdot \dfrac{2y}{2\sqrt{y^2 + \text{R}^2}}}{y^2 + \text{R}^2}\right\} \\ &= \frac{2\pi i\text{R}^2}{(y^2 + \text{R}^2)^{\frac{3}{2}}} \\ &= \frac{2\pi i\text{R}^2}{x^3} \text{ oersted.}^{1} \end{aligned}$$

Intensity on the Axis of a Solenoid

Let AA′ and BB′ be the ends of the solenoid (Fig. 173). Let P be a point on the axis where the intensity is to be found.

[1] At the centre $X = \text{R}$, and the result of p. 128 is obtained.

Let the solenoid of radius R cm. carry n turns of wire per cm. length, the current in the wire being i abamps. Let us consider an element of length of the solenoid dy long, distant y from P. Let the plane AA′ be y_1 from P, and the plane BB′ y_2 from P. Fig. 174 shows this in simplified form.

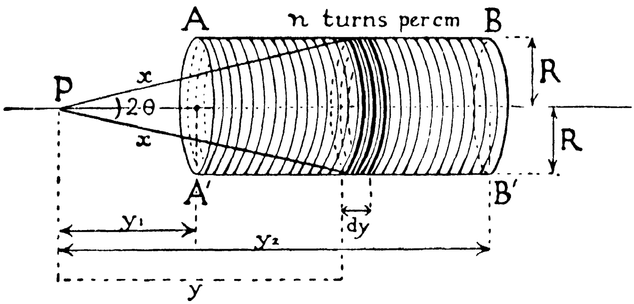


FIG. 173.

The element of the solenoid dy long, along the axis, may be regarded as a circular coil carrying a current $nidy$.

Let the diameter of the coil subtend an angle 2θ at P, and let the circumference of the coil be distant x from P, so that $x^2 = y^2 + R^2$.

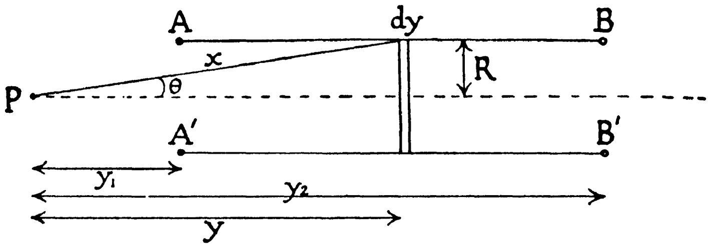


FIG. 174.

Now since the current flowing round our dy-element is $nidy$, the intensity at P due to it is

$$\frac{2\pi niR^2dy}{x^3}.$$

Clearly $\qquad dy = xd\theta \operatorname{cosec} \theta$ (Fig. 175),

and $\qquad \dfrac{R}{x} = \sin \theta.$

So $\qquad \dfrac{2\pi ni R^2 dy}{x^3} = 2\pi ni \sin \theta \, d\theta.$

So the whole field at P is given by

$$\begin{aligned} H &= 2\pi ni \int_{y=y_1}^{y=y_2} \sin \theta \, d\theta \\ &= 2\pi ni \Big[- \cos \theta \Big]_{y=y_1}^{y=y_2} \\ &= 2\pi ni(\cos \theta_1 - \cos \theta_2), \end{aligned}$$

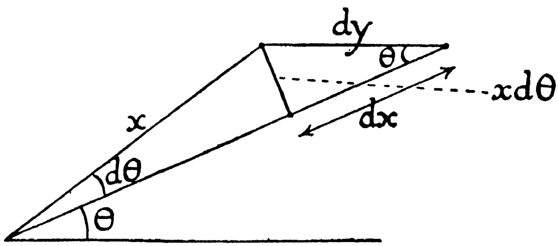


FIG. 175.

if θ_1 and θ_2 are the values of θ when $y = y_1$ and y_2 respectively.

Intensity on the Axis of a Long Solenoid

For ordinary practical purposes most solenoids count as "long." That is, the intensity inside them is the same as if their ends were infinitely far away.

In this case, if $\qquad \theta_1 = 0, \theta_2 = 180°.$

So $\qquad \cos \theta_1 = 1$, and $\cos \theta_2 = -1,$

hence $\qquad \begin{aligned} H &= 2\pi ni\{1 - (-1)\} \\ &= 4\pi ni. \end{aligned}$

A most important and useful result.

Work done in taking Unit Pole round a Current

This conception is of the greatest value in electromagnetic theory. The reason for its importance will appear in Chapter V.

Suppose the plane of a circuit carrying the current is perpendicular to the plane of the paper. Let A and B be points where the wire plunges through the paper in Fig. 176. Let the line joining A to B be the line where the magnetic shell equivalent to the coil cuts the paper.

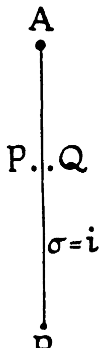


FIG. 176.

The strength of this shell is σ, which is equal to i, the current.

Consider now the difference of magnetic potential between a point P right up against one side of the shell, and a point Q right up against the other side.

The shell subtends a solid angle of 2π both at P and at Q. So the potential at P is $2\pi i$ and at Q is $-2\pi i$.

So the potential difference, which is numerically equal to the work done in taking unit pole from P to Q by any path (*e.g.* round the outside of the wire carrying the current) is

$$2\pi i - (-2\pi i),$$

or

$$4\pi i \text{ ergs.}$$ [1]

Line Integral of a Magnetic Field

The electromotive force in a circuit is the work done in taking unit charge round the circuit against the electric field. It may thus be called the "Line Integral of the Electric Field."

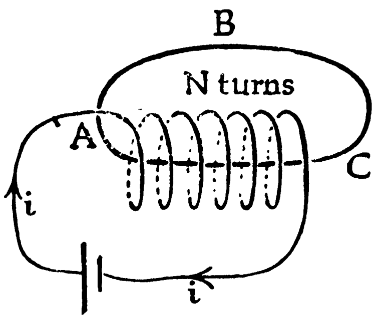


FIG 177.

In the same way the work done in taking unit pole round a closed path, against the magnetic field, is known as the Line Integral of the Magnetic Field. It is also called the Magneto-motive Force, and its use in this connection will be explained in Chapter V.

Ampère's Circuital Theorem states that the Line

[1] You may prefer to prove this by allowing P and Q to approach the shell from opposite sides.

Integral of a Magnetic Field is $4\pi \times$ the total current encircled. Thus, in a solenoid with N turns altogether, the work done in taking unit pole round a closed circuit linked with all the turns, such as the circuit ABC in Fig. 177, will be $4\pi Ni$. This follows at once from the last section, since the work done in threading each turn is $4\pi i$, and there are N turns. Similarly, if an endless circuit of N turns is wound on an anchor-ring, the work done in taking unit pole along any closed path, right round the anchor-ring inside the windings, will be again $4\pi Ni$ if i is the current carried. This result is extensively used in electromagnetic problems. It is further explained in Chapter V.

Intensity near a Long Straight Current

$4\pi i$ is the work done in taking unit pole by any path round the current. Consider a circular path of radius R round the wire. Let the field be H. The distance covered against $H = 2\pi R$.

Then
$$2\pi RH = 4\pi i.$$
$$\therefore \quad H = \frac{2i}{R}.$$

Couple acting on a Coil in a Magnetic Field

A coil of area A carrying a current i is equivalent, by Ampère's Theorem, to a magnet of moment Ai. If the plane of the coil makes an angle θ with the field, the axis of the equivalent magnet makes a complementary angle. It is easily seen that the couple is $H \cos \theta \times$ the magnetic moment. So the couple on the coil is $HAi \cos \theta$.

CHAPTER III

ELECTROMAGNETISM BY CURRENT-ELEMENT METHOD

Unit Current—Intensity on the Axis of a Circular Coil—Intensity near a Long Straight Current—Work done in taking Unit Pole round a Current—Force on a Current in a Magnetic Field—Couple on a Coil in a Magnetic Field—Force between two Straight Currents—Alternative Definition of Electromagnetic Unit of Charge.

Unit Current

An alternative way of defining current in terms of magnetic field, which can be shown to lead in all cases to the same result, is as follows:

The electromagnetic unit of a current is of such a size that a current i, flowing in an element of wire of length dx long,

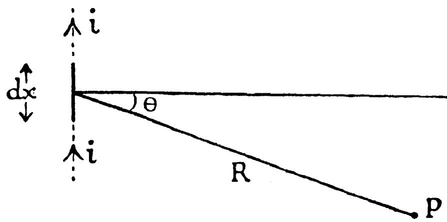


Fig. 178.

produces at a point P, distant R away in a direction making an angle θ with the normal to dx, a field

$$\frac{idx \cos \theta}{R^2} \text{ oersted}$$

in a direction perpendicular to R and to dx (see Fig. 178).

This has a very simple application in the Tangent Galvanometer, where θ is the angle at the centre and R is constant, for

$$H = \int_0^{2\pi Rn} \left(\frac{i}{R^2}\right) dx = \frac{2\pi ni}{R} \text{ oersted.}$$

This definition thus leads directly to the ordinary elementary definition of the absolute e.m. unit as "the current that, flowing in 1 cm. of wire bent into the arc of a circle of 1 cm. radius, produces an intensity of 1 oersted at the centre."

Intensity on the Axis of a Circular Coil

Consider the same coil as in Fig. 172 in the last chapter.

A current-element dz long, at A on the coil, causes at P an intensity $\frac{idz}{x^2}$ in a direction perpendicular to AP.

Since by symmetry the total intensity at P acts along the axis, we need only consider the component of $\frac{idz}{x^2}$ which is effective along the axis.

This component is $$\frac{idz}{x^2} \sin \text{DPA},$$

or $$\frac{idz}{x^2} \cdot \frac{\text{R}}{x}.$$

∴ Total intensity is the sum of all such components caused by elements all round the circle of the coil.

Thus $$\text{H} = \int_0^{2\pi\text{R}} \frac{i\text{R}}{x^3} dz$$
$$= \frac{2\pi i\text{R}^2}{x^3} \text{ oersted.}[1]$$

This was the result of the last chapter. From this result, as before, the field along the axis of a solenoid can be found.

Intensity near a Long Straight Current

Let AB be part of an infinitely long straight wire, carrying a current i. Let P be a point distant r from the wire, and let PN be perpendicular to AB (Fig. 179).

Consider an element of the wire at Q, dy long, distant y from N. Let its distance from P be x, so that $x^2 = y^2 + r^2$.

[1] At the centre $x = \text{R}$, and the result of p. 128 is obtained.

The intensity at P due to dy

$$= \frac{i dy \cos \mathrm{NPQ}}{x^2}$$

$$= \frac{irdy}{x^3}$$

$$= \frac{irdy}{(r^2 + y^2)^{\frac{3}{2}}}.$$

Now this intensity is perpendicular to the plane of the paper.

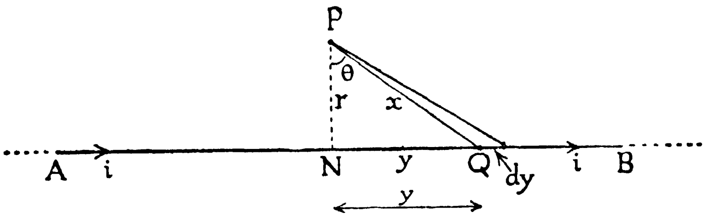


FIG. 179.

So the intensities due to all the elements along the wire add up directly. So at P we have

$$\mathrm{H} = \int_{-\infty}^{+\infty} \frac{irdy}{(r^2 + y^2)^{\frac{3}{2}}}$$

if $$\frac{y}{r} = \tan \theta$$

$$y = r \tan \theta$$

$$dy = r \sec^2\theta \, d\theta$$

$$(r^2 + y^2)^{\frac{3}{2}} = r^3 \sec^3\theta.$$

So we have

$$\mathrm{H} = \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{i \cos \theta \, d\theta}{r}$$

$$= \frac{2i}{r} \text{ oersted}$$

Work done in taking Unit Pole round a Current

From the last result, the work done in taking unit pole round a current at a distance r from it may be found. The pole has to go a distance of $2\pi r$ against a force of $\frac{2i}{r}$ dynes. So $2\pi r \times \frac{2i}{r}$ ergs of work are done. Thus work done $= 4\pi i$ ergs, and is the same for all distances from the wire.

Force on a Current in a Magnetic Field

Though this method is, in general, clumsy for determining fields, except when their components are all in the same direction, it is very useful for determining forces on currents

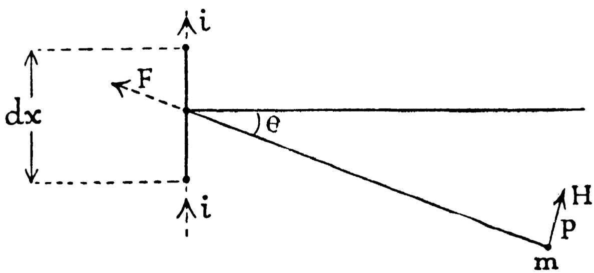


FIG. 180.

in magnetic fields, between poles and currents, and between adjacent currents.

If a pole m be placed at P in Fig. 180, the force on it, if dH is the field due to the current-element dx,

$$= m\,dH, \text{ downwards into the paper.}$$

$$= \frac{midx \cos\theta}{R^2}.$$

By Newton's Third Law there is an equal and opposite force on the current-element due to the pole m.

But from the point of view of the current-element, this force is caused by the field $\frac{m}{R^2}$ set up by the pole m.

Thus the force on the current-element due to a field F, in the direction shown, is

$$Fidx \cos \theta$$

in a direction perpendicular to F and to dx.

Usually we only consider cases where F is perpendicular to dx. In such cases the sense of the force is given by Fleming's Left-hand Rule, as explained in Chapter XIII, Part I, p. 230.

Thus the force in dynes on a straight conductor of length l carrying a current of i absolute e.m. units, in a field of H oersted, perpendicular to the conductor, *in vacuo*, is

$$Hil \text{ dynes.}$$

Note the reservation *in vacuo*. In a medium of permeability μ the force is μHil, or Bil dynes, as is shown at the end of Chapter IX, Part II, p. 477.

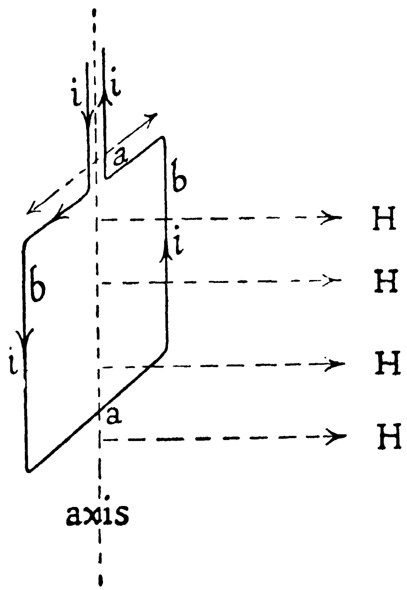


FIG. 181.

Thus in the general case the force on a conductor carrying a current in a magnetic field is

$$i\int \mu H \cos \theta dl$$

where θ has the meaning assigned to it in Fig. 180.

Couple on a Coil in a Magnetic Field

If the coil shown in Fig. 181 makes an angle θ with the field *in vacuo*, the force on each vertical side $= biH$. In the figure, by Fleming's Left-hand Rule, this force is turning the coil in an anticlockwise direction, looked at from above.

Taking moments about the vertical axis of the coil

$$\text{Moment of Couple} = bi\text{H} \cdot \frac{a}{2} \cos\theta + bi\text{H} \cdot \frac{a}{2} \cos\theta$$
$$= abi\text{H} \cos\theta$$
$$= \text{A}i\text{H} \cos\theta$$

where A is the area of the coil.

It is easy to show that this is true whatever the shape. The coil need not be rectangular as it is in the figure.

The couple on the horizontal sides is zero, because the force acts vertically upwards or downwards. Thus—

$$\text{Couple} = \text{Field} \times \text{area} \times \text{current} \times \cos\theta$$
$$\text{C} = \text{HA}i \cos\theta.$$

It is zero when the plane of the coil is perpendicular to the field, and AiH, a maximum, when the plane of the coil is parallel to the field.

Force between two Straight Currents

If two straight currents, i_1 and i_2, run parallel R cm. apart *in vacuo*, the field due to i_1 at any point on i_2 is, as we have seen,

$$\frac{2i_1}{\text{R}} \text{ oersted.}$$

So in a length x of one current (the other being infinite in length), provided x is great compared with R, the total force of attraction if the currents have the same sense (or repulsion if they have different senses) is

$$\frac{2i_1i_2x}{\text{R}} \text{ dynes.}$$

In a medium of permeability μ this force is

$$\frac{2\mu i_1i_2x}{\text{R}} \text{ dynes,}$$

for a reason which is explained at the end of Chapter IX, Part II, p. 477.

Alternative Definition of Electromagnetic Unit of Charge

We have seen that the force on 1 cm. of wire carrying a current at right angles to a magnetic field is Hi. If we regard the current as a charge e moving with velocity v, then $i = ev$.

So the force F on a charge e moving across a magnetic field with velocity v is given by

$$F = Hev.$$

Thus if H is 1 oersted, v 1 cm./sec., and F 1 dyne, e must be one electromagnetic unit of charge.

Thus the electromagnetic unit of charge may be defined as the charge which experiences a force of 1 dyne when it is moving at 1 cm./sec. perpendicular to a magnetic field of 1 oersted, *in vacuo*.

CHAPTER IV

ELECTROMAGNETIC INDUCTION

Fundamental Equations—Measurement of Flux or Field-intensity—The Earth-inductor—Average Current in a Rotating Coil—Faraday's Disc—Self-inductance and Mutual Inductance—Energy stored in an Inductive Circuit—Rate of Growth and Decay of Currents—E.M.F.s in Free Space—The Electrodeless Discharge.

Fundamental Equations

THE E.M.F. round a closed circuit of one turn is the rate of change of flux (total number of lines) threading the circuit. The unit of flux is the line, or Maxwell.

Since most coils have more than one turn, we shall define N in our equations as the number of line-linkages, or flux $\times$ number of turns. We shall say that N is measured in Maxwell-turns.

Suppose we consider the positive direction of the E.M.F. to be the direction of flow of the current which sets up the flux we are considering. By Lenz's Law, the induced E.M.F. opposes this current if the flux increases. So for the E.M.F. at any instant we may write

$$E = -\frac{dN}{dt} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

If the resistance of the circuit is R, the current $i = \frac{E}{R}$,

so
$$i = -\frac{1}{R}\frac{dN}{dt} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

provided that R does not approach zero, and that the circuit has no inductance. The effect of inductance is treated later.

If Q is the total charge passing a point in the circuit in time t,

$$\begin{aligned} Q &= \int_0^t i\,dt \\ &= \int_0^t -\frac{1}{R}\frac{dN}{dt}dt \\ &= -\frac{1}{R}\int_0^t dN \\ &= -\frac{N}{R}. \qquad \qquad (3) \end{aligned}$$

Thus the charge passing is simply the total change of Maxwell-turns divided by the resistance. The negative sign is only concerned with the direction of flow of the charge, and need not as a rule be taken into account. The equations for E, i, and Q tell us all we need about Electromagnetic Induction. All other equations dealing with changing currents are built on these three.

Measurement of Flux or Field-intensity

Equation 3 can be used directly to measure such quantities as the total flux in the air-gap of an electromagnet, or the horizontal, vertical, or total strength of the earth's field. Both a Ballistic Galvanometer and a Fluxmeter are instruments to measure the total charge passing. They are discussed in great detail in Part II, Chapter VI, pp. 389–409. They need only be considered here as enabling us to find the total quantity of charge passing.

The principle of the measurement is to find the total change, in Maxwell-turns, for a coil of known area and number of turns, when it is either removed suddenly from an air-gap or turned through a known angle.

Let us suppose, for example, that a search-coil of 100 turns entirely surrounds the flux of the air-gap of a big electromagnet. Of course if the total flux goes through the coil, and we are measuring flux, not field-intensity, we need not know the area of the coil.

Suppose the total charge passing, when the coil is removed from the air-gap, is 4 microcoulombs, and the resistance of the coil-galvanometer circuit is 20 ohms.

Since $Q = -\frac{N}{R}$, we have

$$N = QR$$
$$= (4 \times 10^{-6}) \times 10^{-1} \times 20 \times 10^{9}$$

(since 1 coulomb $= 10^{-1}$ e.m. units and 1 ohm $= 10^{9}$ e.m. units).

$$\therefore \quad N = 8000 \text{ Maxwell-turns,}$$

$$\therefore \quad \text{Flux} = \frac{8000}{100} = 80 \text{ Maxwells.}$$

Suppose we had wanted the intensity rather than the flux, and had obtained the same results with a search-coil of average area-turns 10 sq. cm. Clearly the average field over the area covered by the coil, before it was removed, would have been given by

$$H = \frac{80}{10} = 8 \text{ oersted.}$$

The Earth-inductor

A large flat coil, called an Earth-inductor, provides, with a ballistic galvanometer, or fluxmeter, the quickest way of measuring the horizontal, vertical, or total intensity of the earth's magnetic field. Let us first measure the vertical component. Suppose such a coil, having n turns of average area a sq. cms., is laid on a horizontal table, and connected in series with a ballistic galvanometer or fluxmeter. The total resistance of the circuit is r ohms. When the coil is turned over and laid on its other side (being thus turned through 180°) a charge q passes through the galvanometer.

As the coil turned through 90°, the flux through it changed from Va to zero, and as it turned through the second 90° the flux changed further from zero to $-Va$. So the total change of flux was $2Va$. [V is the vertical component of the earth's field.] Thus the change of Maxwell-turns was $2Van$. So by our equation 3,

$$q = \frac{2Van}{r}.$$

We must, however, express q and r in absolute units, and have probably measured them in microcoulombs and ohms respectively. Now 1 microcoulomb $= 10^{-7}$ e.m. units, and 1 ohm $= 10^{9}$ e.m. units.

So if we express q in microcoulombs and r in ohms,

$$10^{-7} \times q = \frac{2Van}{10^9 \times r}$$

so
$$V = \frac{50qr}{an}.$$

A fairly ordinary kind of ballistic galvanometer has a resistance of 25 ohms, and a sensitivity of 100 mm. per microcoulomb. Suppose that with such a galvanometer the throw was 220 mm. with an earth-inductor of 100 turns of average area 75 sq. cms., having a total resistance of 5 ohms—

$$V = \frac{50 \times 2{\cdot}2 \times 30}{75 \times 100}$$
$$= 0{\cdot}44 \text{ oersted.}$$

H can be determined by turning the coil through 180° about a vertical axis from a position in which its normal points along the magnetic meridian, and the dip can be found from $\tan^{-1}\frac{V}{H}$. I can be found as $\sqrt{V^2 + H^2}$, or by turning the coil through 180° from a position with its plane perpendicular to the direction of I.

Average Current in a Rotating Coil

If a coil (Fig. 182) having its axis PQ along a magnetic field H and its plane (at the start) perpendicular to the field, is rotated about any diameter AB, it is clear that the flux

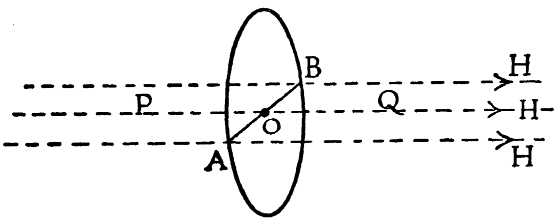


FIG. 182.

will keep on changing as it did with the earth-inductor. The direction of flow of the current will also be the same throughout the first half-turn. Throughout the next half-turn, however,

it will be reversed; and if there is no commutator in the circuit the average current will be zero. If, however, there is a commutator in the circuit, so that all the charges circulate in the same direction, the total current is easily calculated.

Suppose a coil of n turns, of area a, is rotating about a diameter in a magnetic field H, with a frequency of ν (that is, making ν rotations per second). The resistance of the circuit through which it sends the current is r. There is assumed to be a commutator in the circuit. Then by the last section, on earth-inductors, the charge q flowing in one half-turn is

$$q = \frac{2\text{H}an}{r}.$$

There are 2ν half-turns per second, and i is the charge per second, so

$$\begin{aligned} i &= 2\nu\text{Q} \\ &= \frac{4\text{H}an\nu}{r}. \end{aligned}$$

If i is in amperes and r in ohms, we must re-write this

$$\frac{i}{10} = \frac{4\text{H}an\nu}{10^9 r}.$$

Thus if i is in amperes and r in ohms,

$$i = \frac{4\text{H}an\nu}{10^8 r}.$$

As the E.M.F. is $i \times r$, we may say the average E.M.F. during a half-cycle in the coil is given by

$$\text{E} = \frac{4\text{H}an\nu}{10^8}.$$

(The alternating average is, of course, zero.)

Thus in a coil of 50 turns, having area 10 sq. cms. and resistance 5 ohms, revolving 100 times a second in a field of 80 gauss, the average current is given by

$$\begin{aligned} i &= \frac{4 \times 80 \times 10 \times 50 \times 100}{10^8 \times 5} \\ &= 0{\cdot}016 \text{ ampere.} \end{aligned}$$

It should be noted that this "average current" is *not* the

root–mean–square current measured by a thermo-ammeter. This average current is $\frac{2}{\pi}$ of the maximum current, whereas the R.M.S. current is $\frac{1}{\sqrt{2}}$ of the maximum. Thus the average current, from the point of view of passage of charge, is $\frac{2\sqrt{2}}{\pi}$ of the average current from the point of view of heating effect. The mean value of $\sin\theta$ over a half-cycle is $\frac{2}{\pi}$, whereas the R.M.S. value of $\sin\theta$ is $\frac{1}{\sqrt{2}}$.

The average E.M.F. is thus

$$E = 0{\cdot}080 \text{ volt}$$

(These results only apply when a commutator is used.)

Faraday's Disc

This apparatus consists simply of a disc of conducting material (such as copper) which rotates in a plane perpendicular to a magnetic field. A brush bears on the circumference. An E.M.F. is generated between the axle and circumference because the moving conductor cuts the lines of force.

Let H be the field, r the radius, and n the number of revolutions of the disc per second. In time dt the disc turns through an angle of $2\pi n dt$ radians (since it does n complete turns in one second). Any one radius thus sweeps out an area of $\pi n r^2 dt$ sq. cms. (because the area of a very thin sector subtending an angle $\delta\theta$ is $\frac{1}{2}$ base $\times$ height, or $\frac{1}{2}\, r\delta\theta \times r$, or $\frac{1}{2}\, r^2\delta\theta$).

The flux cut in time dt is thus $H\pi n r^2 dt$ Maxwells.

Thus the rate of change of flux is $\pi n r^2 H$, and this is the E.M.F. in absolute units. In volts the E.M.F. is $10^{-8}\pi n r^2 H$. Thus a disc of radius 10 cm. revolving 100 times per second in a field of 2000 gauss would generate an E.M.F. of

$$10^{-8} \times \pi \times 100 \times 10^2 \times 2000 \text{ or } 0{\cdot}628 \text{ volt}$$

Self-inductance and Mutual Inductance (See p. 248.)

L is the usual symbol for self-inductance; M for mutual inductance.

Thus if the current i is changing, an E.M.F. of $\left(-\mathrm{L}\frac{di}{dt}\right)$ acts round the circuit in the direction of flow of the current. If the mutual inductance between two circuits carrying currents i_1 and i_2 is M, then the induced E.M.F. in the i_1 circuit is $\left(-\mathrm{M}\frac{di_2}{dt}\right)$, and in the i_2 circuit it is $\left(-\mathrm{M}\frac{di_1}{dt}\right)$.

The calculation of inductance is mathematically very difficult. An attempt to calculate the magnetic intensity, at a point distant x from the centre of a single coil of wire of radius r, will immediately show why the calculation is difficult. The intensity H will be found to be given by

$$\mathrm{H} = 2ir\int_0^\pi \frac{(r - x\cos\theta)d\theta}{(x^2 + r^2 - 2rx\cos\theta)^{\frac{3}{2}}},$$

which is not very promising. To find the total flux through the coil we should have to perform this integration and then another of H over the coil. When we had solved this problem we should have done only the simplest possible case. Actual inductance problems are solved by special methods.

A very rough approximation for the inductance of a single coil may be made by assuming the intensity is the same all over as it is at the centre. The intensity at the centre due to unit current in the coil is, as we have seen, $\frac{2\pi}{r}$ oersted. Thus since the area is πr^2, the self-inductance is very roughly $2\pi^2 r$. This is also the mutual inductance of two single turns wound together.

If a fairly long solenoid has n turns of wire in a length l, we may assume that the intensity inside it is $\frac{4\pi n}{l}$ oersted for unit current (though of course this is only true on the axis, and for an infinitely long solenoid).

The flux threading it is thus $\frac{4\pi^2 nr^2}{l}$, and the number of Maxwell-turns is $\frac{4\pi^2 n^2 r^2}{l}$. This is the self-inductance; and similarly the mutual inductance for two solenoids, wound on top of one another closely, is $\frac{4\pi^2 n_1 n_2 r^2}{l}$.

This formula depends upon the further (and even more

inaccurate) assumption that all the flux is linked with all the turns. It may thus easily be 20% too high.

We have considered absolute units only. Now inductance is measured in E.M.F. per rate of change of current. Thus the inductance in volts per amp.-per-second will be $\frac{10^{-1}}{10^8}$, or 10^{-9} of its value in abvolts per abamp.-per-second.

Thus the numerical value of the inductance in henries is 10^{-9} of its value in absolute units, or

$$1 \text{ Henry} = 10^9 \text{ abhenries.}$$

Similarly the inductance of a solenoid in henries is

$$L = \frac{4\pi^2 n^2 r^2}{10^9 l}.$$

If the solenoid contains iron the flux is enormously increased; in the ratio of the permeability if the iron fills the whole solenoid. It may thus be multiplied 1000 times easily.

Energy stored in an Inductive Circuit

A growing current in an inductive circuit must be forced along against a back E.M.F. $L\frac{di}{dt}$. The battery, therefore, must do more work than if the circuit was non-inductive. In time dt a charge of idt units must pass. Thus the work done in time $dt = L\frac{di}{dt} \times idt$, or $Lidi$ ergs.

Thus the total work, integrating over all the time needed to supply a steady current i_0

$$= \int_0^{i_0} Lidi = \tfrac{1}{2}Li^2 \text{ ergs.}$$

This work is stored as energy in the circuit, and reappears mainly in the heat of the inductive spark at break. If L is in henries and i in amperes, it is easily seen that the energy is given in joules.

Thus a typical "smoothing" coil for direct-current mains used for high-tension of a wireless set, of inductance 20 henries carrying 100 milliamps. or 0·1 amp., the energy stored is $\frac{1}{2} \times 20 \times 0{\cdot}01$ joules or 0·1 joule.

Rate of Growth and Decay of Currents

Imagine a steady E.M.F. E is suddenly applied to a circuit of resistance R and self-inductance L. The current i as it grows will produce a back E.M.F. of $L\frac{di}{dt}$.

Thus we shall have, considering the whole circuit,

$$\text{Current} = \frac{\text{Effective E.M.F.}}{\text{Resistance}}$$

$$i = \frac{E - L\frac{di}{dt}}{R}.$$

This equation is more usually written as expressing the fact that E is equal to the opposing E.M.F. + the ohmic potential drop.

$$E = Ri + L\frac{di}{dt}.$$

Thus $$dt = \frac{Ldi}{E - Ri}$$

$$t = L\int\frac{di}{E - Ri}$$

$$= -\frac{L}{R}\log(E - Ri) + \text{constant}.$$

By hypothesis $i = o$ when $t = o$.

So $$o = -\frac{L}{R}\log E + \text{constant}.$$

$$\text{Constant} = \frac{L}{R}\log E.$$

$$\therefore \quad t = \frac{L}{R}[\log E - \log(E - Ri)]$$

$$= \frac{L}{R}\log\frac{E}{E - Ri}$$

$$= \frac{L}{R}\log\frac{i_0}{i_0 - i}$$

where $i_0\left(=\frac{E}{R}\right)$ is the final steady current.

Writing the equation exponentially and turning it upside down,

$$\frac{i_0 - i}{i_0} = e^{-\frac{R}{L}t}$$

Thus $$i = i_0\left[1 - e^{-\frac{R}{L}t}\right].$$

The quantity $\frac{L}{R}$ is usually called the Time-constant of the circuit, since the rate at which the current builds up depends on it alone.

When
$$t = \frac{L}{R},$$
$$\begin{aligned} i &= i_0(1 - e^{-1}) \\ &= i_0\left(1 - \frac{1}{2\cdot718\ldots}\right) \\ &= 0\cdot632 i_0. \end{aligned}$$

The time-constant is thus the time it takes for a current to reach 0·632 of its final value.

The ratio is true both for absolute units and for henries and ohms, since, of the henry and ohm, each is 10^9 of its corresponding absolute unit.

Thus a circuit, in which the numerical values of inductance and resistance coincide, reaches 0·632 of its final current-value in one second.

When E is shut off and the current decays, the equation obviously remains unchanged except that $E = O$.

So
$$L\frac{di}{dt} + Ri = O.$$

By the same mathematical process, merely putting $E = O$ in the result
$$i = i_0 e^{-\frac{Rt}{L}}.$$

When $Rt = L$ in this circuit, $i = \frac{i_0}{2\cdot718\ldots}$.

So a current which reaches 0·632 of its full value in one second when starting up reaches $\frac{1}{2\cdot718}$ of its full value in one second when collapsing.

E.M.F. in Free Space. The Electrodeless Discharge

We have been considering the magnetic fields set up by moving charges; or rather (for it comes to the same thing) by moving electric fields. It is important to realize that electric fields are set up in the same way by moving magnetic

fields, for this realization was the foundation of Maxwell's work on the Electromagnetic Theory of Light, and is necessary for our understanding of this theory.

A moving electric field sets up a magnetic field at right angles to it. Similarly a moving magnetic field sets up an electric field at right angles to it.

Suppose in Fig. 183 an alternating current is fed to the coil C from the terminals AB. The magnetic lines of force through C due to this current will run in a direction perpendicular to the plane of the coil, and will alternate in sense. They should thus be encircled by a ring of alternating E.M.F. in the plane of the coil, just as an alternating current along the axis of the coil would produce an alternating magnet field encircling it in the plane of the coil. If a sphere D containing neon, which emits light under the action of a field so low as 30 volts/cm., is placed inside the coil C, a circular luminous discharge is seen in it in the path of the alternating E.M.F. if the magnitude and frequency of the original alternating current in C are big enough.

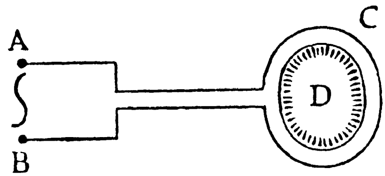


Fig. 183.

The discharge in the neon is thus electrodeless, and is due to an E.M.F. in free space.

This experiment illustrates directly the fact which Maxwell grasped when he worked out the Electromagnetic Theory of Light. Practically, it led to modern optical theory and to wireless telegraphy. Theoretically it led to much of our modern view of the nature of the physical universe. It must form the basis of any Unified Field theory which relates Gravitation and Electromagnetism. The Electric and Magnetic Fields are complementary. It is common knowledge that a moving electric field is linked with a magnetic field; but it is not always realized that a moving magnetic field is linked in exactly the same way with an electric field. If solids are present they may carry currents or be magnetized. Solids are, however, not necessary; the whole thing may happen in free space.

*

CHAPTER V

MAGNETIC MATERIALS

H, B, μ, I, S—Relations between H, B, μ, I, and S—Paramagnetism, Diamagnetism, and Ferromagnetism—Theories of Diamagnetism, Paramagnetism, and Ferromagnetism—The Magnetic Circuit—Anchor-rings with or without Air-gaps—Energy per Unit Volume of Magnetic Field—Force in the Air-gap of an Electromagnet—Hysteresis—Measurement of Hysteresis: Direct Method—Measurement of Hysteresis: Ballistic Method.

H, B, μ, I, S

H, the magnetic intensity, corresponds to E, the electric intensity. Its value at any particular point is numerically equal to the number of dynes of force which would be exerted on a unit pole situated at the point if the presence of this pole did not disturb the field.

The number of lines of force [1] per square centimetre *in vacuo* is chosen to be equal to H. A unit magnetic intensity, having one line per square centimetre, is an intensity of 1 oersted.

B, the magnetic induction, corresponds to D, the electric induction or displacement, except that D is numerically equal to $\frac{1}{4\pi}$ of the number of lines of force per square centimetre in the medium (or in other words equal to the number of Faraday tubes, each containing the 4π lines of force which emanate from a unit charge). B, on the other hand, is numerically equal to the actual number of lines of force. When the lines of force occur inside a medium they are called "lines of induction," but every line of induction becomes a line of force at the point where it leaves the medium and enters a vacuum. B is measured in gauss.

The trouble in distinguishing B and H arises from the fact that the force on a pole inside a medium is determined by H, the number of lines that would be there if the medium were

[1] See footnote, p. 69.

replaced by a vacuum, whereas B determines the number of lines that actually leave the surface of the medium at a boundary. The simplest way to visualize the difference

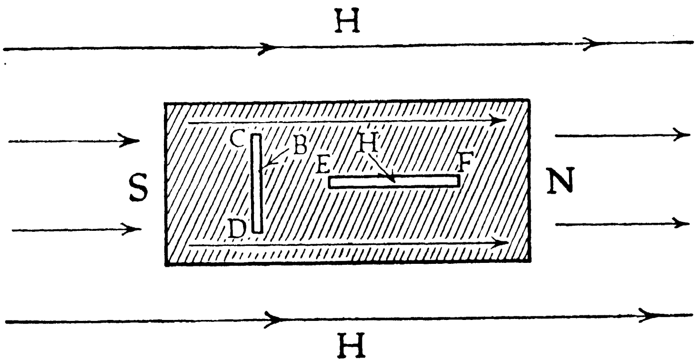


FIG. 184.

between them is to consider two very narrow slits CD and EF cut in a piece of magnetized material SN, magnetized from S to N (Fig. 184).

The whole piece of magnetic material is immersed in a magnetic field of intensity H.

The slit CD is cut perpendicular to the direction of magnetization. Since it is very narrow the number of lines of force crossing it is the same per square cm. as the number in the iron. The force on a unit pole in this slit would thus be B dynes if the slit were indefinitely narrow; or rather the force would approach B dynes as the width of the slit approached zero.

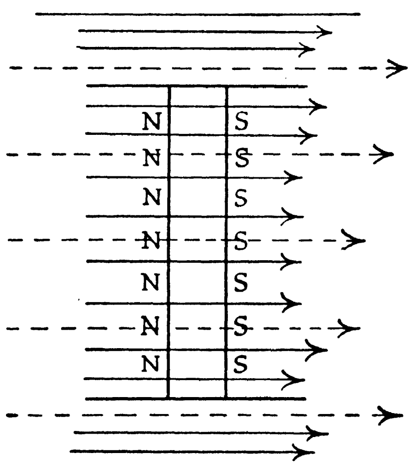


FIG. 185.

We may regard the extra lines of force additional to those due to H as being due to the opposite induced poles on the two sides of the slit. These produce their own lines. This arrangement is shown in Fig. 185. The dotted lines of force are due to H. The continuous lines are due to the magnetized state of the iron except in the slit, where they are caused by the induced poles. Both dotted and continuous lines are included in B.

The slit EF is cut parallel to the direction of magnetization. As its width approaches zero the induced poles on its end approach zero also. Hence only the lines of force due to H are left inside it.

The force on a unit pole in this slit would approach H dynes as the width of the slit approached zero. This arrangement is shown in Fig. 186.

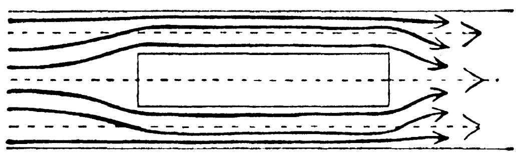


FIG. 186.

μ, *the permeability*, corresponds to k, the dielectric constant. It is the ratio of the induction to the intensity of the magnetizing field, or $\frac{B}{H}$.

It follows from these definitions that since F, the force on a pole m with a field H in air, is given by

$$F = Hm \text{ dynes,}$$

the force on a pole m in a medium of permeability μ with induction B must be given by

$$F = \frac{Bm}{\mu} \text{ dynes.}$$

Also we know that the force between poles m_1 and m_2, situated r cm. apart in air, is

$$F = \frac{m_1 m_2}{r^2} \text{ dynes (by the Inverse Square Law),}$$

so that the force between poles m_1 and m_2, situated r cm. apart in a medium of permeability μ is

$$F = \frac{m_1 m_2}{\mu r^2} \text{ dynes.}$$

I, the Intensity of magnetization, is the amount of pole per unit area of cross-section taken at the end of a magnet perpendicular to the direction of magnetization. It is also the magnetic moment per unit volume anywhere inside the magnet.

This is easily seen to be an equivalent definition.

Let our magnet be of length l and cross-section a, and have pole-strength m.

Then from the first definition $I = \frac{m}{a}$.

The volume is la, and the magnetic moment ml. So from the second definition

$$I = \frac{ml}{la} = \frac{m}{a}.$$

The *susceptibility*, s, is the ratio of the intensity of magnetization to the magnetizing field. It is thus given by

$$s = \frac{I}{H}.$$

Relations between H, B, μ, I, and S

The intensity 1 cm. from a unit point pole in air is 1 oersted. So 1 line of force must be crossing each sq. cm. of the little sphere of 1 cm. radius surrounding the point pole. The area of this little sphere is 4π sq. cm. So 4π lines of force in air, or lines of induction in a magnetic medium, leave each unit pole.

Let us consider now how many lines of induction there are inside a piece of material magnetized to intensity I in a field H.

H lines per sq. cm. penetrate the material because of the field. $4\pi I$ lines per sq. cm. penetrate the material because of the intensity I units of pole per sq. cm. of cross-section of the surface of the disc-like slot shown in Fig. 185. Thus $H + 4\pi I$ lines of induction penetrate the material altogether, so

$$B = H + 4\pi I.$$

If we divide this equation through by H, we get

$$\frac{B}{H} = 1 + 4\pi\frac{I}{H}.$$

Since $\frac{B}{H} = \mu$, and $\frac{I}{H} = s$, by definition, we have

$$\mu = 1 + 4\pi s.$$

The easiest use of these equations can be illustrated by an example.

Question 1.—A piece of iron wire of length 100 cm. and cross-section 2 sq. mm. is placed in a magnetic field of intensity 50 oersted in the direction of its length. Its magnetic moment is then found to be 2000 units. Find its permeability and susceptibility, and the magnetic induction through it.

Its intensity of magnetization, I, is the magnetic moment per unit volume. The volume = 100 × 0·02, or 2 c.c.

So $$I = \frac{2000}{2} = 1000 \text{ units.}$$

Since $B = H + 4\pi I$, we have

$$B = 50 + 4000\pi$$
$$= 12616 \text{ gauss.}$$

Since $$\mu = \frac{B}{H}$$

$$\mu = \frac{12616}{50} = 252.$$

Since $$s = \frac{I}{H}$$

$$s = \frac{1000}{50} = 20.$$

Paramagnetism, Diamagnetism, and Ferromagnetism

Magnetic substances are of three kinds—paramagnetic, diamagnetic, and ferromagnetic.

In paramagnetic substances μ is a little greater than 1, being usually of the order of 1·00001.

Manganese is the most paramagnetic element, having μ = 1·00013, and chromium is next with μ = 1·000047.

Molybdenum and osmium have μ = 1·0000005.

In diamagnetic substances μ is a little less than 1, being of the order of 0·99999.

Bismuth, the most diamagnetic substance, has $\mu = 0{\cdot}9999824$. Copper has $\mu = 0{\cdot}9999989$.

For diamagnetic substances μ is practically independent of the magnetizing field, but for paramagnetic substances μ gets less as the field increases.

For diamagnetic substances s is independent of temperature. For paramagnetic substances s varies inversely as the absolute temperature.

Ferromagnetic substances are completely different. The maximum value of μ is about 8000 for mild steel and about 2000 for soft iron. Nickel has a maximum of about 300, and cobalt of about 170. Iron, nickel, and cobalt are the only ferromagnetic elements. μ is always positive for ferromagnetic substances.

Ferromagnetics have a definite temperature at which they suddenly lose their ferromagnetism and become paramagnetic. This temperature is known as the Curie point, and is about 800° C. for iron, 400° C. for nickel, and 1100° C. for cobalt.

Iron, when cooling, has a number of temperatures causing recalescence, at which a red glow appears and heat is given out. The Curie point is one of the temperatures of recalescence. The obvious inference is that at this temperature iron changes its molecular arrangements from a state of more to one of less energy, and therefore gives out energy, and that both states are paramagnetic, while the second is ferromagnetic as well.

Theories of Diamagnetism, Paramagnetism, and Ferromagnetism

Diamagnetism and paramagnetism are accounted for by supposing that when a magnetic field acts on matter it produces two main effects, which are about to be described.

The molecules of the material are supposed to contain electronic charges describing orbits about centres of force.

These moving charges are equivalent to currents flowing in closed paths, and each has therefore a magnetic field along its axis, and tends to behave like a little magnet.

The applied magnetic field causes a change in the motion of the electron in its orbit. By Lenz's Law this change is

such as to keep the total flux through the circuit constant. It must therefore set up a magnetic field opposing the original applied field.

This is the supposed cause of all diamagnetism, and on this theory all substances, without exception, are diamagnetic.

The second main effect is that such substances as have molecules with a definite magnetic moment tend to have these molecules oriented with their magnetic axes in the direction of the intensity.

The intensity, produced by the combined effect of all these orientations, is in the same direction as the originally acting intensity, and the number of lines of force inside the material is thus greater than that outside.

This is supposed to be the cause of paramagnetism. It is supposed that in paramagnetic substances the paramagnetic effect merely outweighs the diamagnetic effect. Some substances which are paramagnetic at low temperatures become diamagnetic at high temperatures, because their paramagnetism decreases with increase of temperature and their diamagnetism remains unchanged.

Ferromagnetism is supposed to be due to the formation of large groups or units of paramagnetic molecules. These groups are supposed to be capable of turning in a magnetic field, but to experience something like a frictional resistance which requires work to be done either to produce or destroy their alignment along the field. This work is lost irrecoverably as heat, and the loss of energy resulting is known as the hysteresis loss. Ferromagnetism is not yet so well explained as diamagnetism and paramagnetism.

The Magnetic Circuit

Let us consider what happens when a piece of magnetic material is magnetized by having a number of turns of wire carrying a current i amperes wound round it.

The following facts are clear :—

The total number of lines of induction inside the magnet is the same as the total number of lines of force joining the poles outside the magnet.

Let us suppose that the total number of lines of induction

inside our magnetized material is N. It may also be called the Magnetic Flux.

If B and H are uniform across any given cross-section inside the metal, then obviously

$$N = BA = \mu HA$$

if A is the area of cross-section. This is true wherever the cross-section is taken, though of course B and H need not be the same at different cross-sections.

If we take a unit pole right round the track of the flux, going always in the direction of the flux, then the work done on the pole is

$$\int_0 H dx$$

when x is a distance along the path taken always parallel to the flux.

But $\int_0 H dx$ is the line-integral of the field.

But by hypothesis the field is due to n turns of wire carrying a current i amps.

The line integral of n turns carrying i amps. has been shown to be

$$\frac{4\pi ni}{10}.$$

So $$\int_0 H dx = \frac{4\pi ni}{10}.$$

At this point we may pause to notice that we are dealing with two quantities closely analogous to quantities appearing in the electric circuit.

The E.M.F. in a circuit is the work done on unit charge in pushing it round the circuit. It thus corresponds to the Line-Integral. So we call the Line-Integral the Magneto-Motive Force or M.M.F.

The essential properties of a current are: (1) that it is the same all round the circuit, (2) that it is caused by the E.M.F.

The Magnetic Flux is the same all the way round a closed circuit, and is caused by the M.M.F. The flux is thus analogous to the current.

We have now

$$N = \mu HA.$$

So
$$H = \frac{N}{\mu A}.$$

So
$$\int_0 H\,dx = \int_0 \frac{N\,dx}{\mu A}$$

$$= N\int \frac{dx}{\mu A},$$

since N is constant.

But since
$$\int_0 H\,dx = \frac{4\pi ni}{10}$$

we may write

$$N\int \frac{dx}{\mu A} = \frac{4\pi ni}{10},$$

so
$$N = \frac{\frac{4\pi ni}{10}}{\int \frac{dx}{\mu A}}.$$

This equation is analogous to

$$\text{Current} = \frac{\text{E.M.F.}}{\text{Resistance}},$$

and for this reason the quantity

$$\int \frac{dx}{\mu A}$$

is sometimes called the " Magnetic Resistance," and the whole magnetic equation may be written

$$\text{Magnetic Flux} = \frac{\text{Magneto-Motive Force}}{\text{Magnetic Resistance}}.$$

It must be remembered, however, that the magnetic equation refers to statical phenomena, though the electrical equation refers to dynamical phenomena. Nothing is moving round in the flux.

If we realize that our magnetic equation is only an analogy, we shall find it extraordinarily useful for solving problems.

Anchor-rings with or without Air-gaps

The case of an ordinary straight bar magnet is intractable by this method, but the case of an anchor-ring magnetized by a known number of turns of wire is beautifully easy, whether the ring is of uniform or varying cross-section and whether it has an air-gap or not.

Let us consider some numerical examples.

1. An anchor-ring has a uniform cross-section of 12 sq. cm. and a mean diameter of 30 cm. Its permeability is 1000 and it is wound with 200 turns of wire carrying 2 amps. What is the flux through it?

Since A and μ are constant, $\int_0 \frac{dx}{\mu A}$ may be written $\frac{1}{\mu A}\int dx$, or $\frac{2\pi r}{\mu A}$, since dx merely goes round the mean circumference of the ring.

Our equation

$$N = \frac{\text{M.M.F.}}{\text{Magnetic Resistance}}$$

thus becomes

$$\begin{aligned} N &= \frac{\frac{4\pi ni}{10}}{\frac{2\pi r}{\mu A}} \\ &= \frac{2\mu A ni}{10r} \\ &= \frac{2 \times 1000 \times 12 \times 200 \times 2}{10 \times 15} \\ &= 64{,}000 \text{ Maxwells.} \end{aligned}$$

2. Now let us insert an air-gap 1 millimetre wide.

$$\int \frac{dx}{\mu A} \text{ becomes } \left(\frac{0{\cdot}1}{1 \times A} + \frac{2\pi \times 15 - 0{\cdot}1}{1000A}\right)$$

which $$= \frac{0{\cdot}19415}{A}.$$

So $$\begin{aligned} N &= \frac{4\pi ni}{10} \Big/ \frac{0{\cdot}19415}{A} \\ &= \frac{4\pi niA}{1{\cdot}9415} \\ &= 31{,}074 \text{ Maxwells.} \end{aligned}$$

Thus the effect of introducing the gap of 1 mm. is the reduction of the flux to less than half its former value.

If the anchor-ring has different cross-sections in different parts, $\int_0 \frac{dx}{\mu A}$ can obviously be found by calculating the magnetic resistance of each bit separately and adding up these resistances to find the total magnetic resistance.

We can easily get an approximate solution (quite near enough for all practical purposes) for a horse-shoe electromagnet with a keeper and a small air-gap.

If the lengths of keeper and cross-piece are x_1, of the uprights x_2, and of each of the two air-gaps (one at each end of the keeper) d, and if the cross-section of the keeper is A_1, of the cross-piece A_2, and of the uprights A_3, then for n turns carrying current i

$$\text{M.M.F.} = \frac{4\pi ni}{10}.$$

$$\text{Magnetic Resistance} = \frac{2d}{A_3} + \mu\left(\frac{x_1}{A_1} + \frac{x_1}{A_2} + \frac{2x_2}{A_3}\right)$$

so

$$N = \frac{4\pi ni}{10} \Big/ \left\{\frac{2d}{A_3} + \mu\left(\frac{x_1}{A_1} + \frac{x_1}{A_2} + \frac{2x_2}{A_3}\right)\right\}.$$

The problem has now become perfectly simple. Sometimes ni is known as the "ampere-turns" because the M.M.F. due to $\frac{1}{2}$ an amp. going round 20 times or 2 amps. going round 5 times is obviously the same as that due to 1 amp. going round 10 times.

Since $\frac{4\pi}{10} = 1{\cdot}257$, it is sometimes said that

$$\text{M.M.F.} = 1{\cdot}257 \text{ (ampere-turns)},$$

a statement which is apt to be obscure if one does not fully understand its origin.

Energy per Unit Volume of Magnetic Field

We have already seen in the last chapter that the magnetic energy of an inductive circuit (or in other words the energy of the magnetic field set up by such a circuit) is $\frac{1}{2}Li^2$.

We can calculate from this the energy per unit volume of a magnetic field.

Imagine an anchor-ring of material of permeability μ. Let its mean radius be r, and the cross-section of the magnetic material of the ring be a. Let it be a large thin ring, so that without serious error we may take its volume as $2\pi ra$. Let it be wound with n turns carrying a current i. Then its magnetomotive force is $4\pi ni$, and its magnetic resistance

$$\int_0 \frac{dx}{\mu a}, \text{ or } \frac{2\pi r}{\mu a}.$$

The flux is thus $$\frac{4\pi ni}{2\pi r/\mu a}, \text{ or } \frac{2\mu ani}{r}.$$

If this is linked n times with the current (because there are n turns) the Maxwell-turns are $\frac{2\mu an^2 i}{r}$ for current i. So the self-inductance $L = \frac{2\mu an^2}{r}$, the Maxwell-turns per unit current.

$$\text{Thus the magnetic energy} = \tfrac{1}{2} \cdot \frac{2\mu an^2}{r} \cdot i^2$$
$$= \frac{\mu an^2 i^2}{r}.$$

$$\text{Now H, the field,} = 4\pi i \times \text{(turns per unit length)}$$
$$= \frac{4\pi ni}{2\pi r} = \frac{2ni}{r}.$$

So $$i = \frac{rH}{2n}.$$

Substituting for i in the equation for the energy,

$$\text{Energy} = \frac{\mu an^2}{r} \cdot \frac{r^2 H^2}{4n^2}$$
$$= \left(\frac{\mu H^2}{8\pi}\right) \times 2\pi ra.$$

But this is the total energy for a volume of $2\pi ra$. So the energy per unit volume $= \frac{\mu H^2}{8\pi}$.

Since $B = \mu H$, we have also

$$\text{Energy} = \frac{B^2}{8\pi\mu}.$$

It is interesting to compare the energy per unit volume of

the magnetic field, $\frac{\mu H^2}{8\pi}$, with that per unit volume of the dielectric for the electric field, $\frac{kE^2}{8\pi}$. The parallel is exact.

Force in the Air-gap of an Electromagnet

We are considering only the force of attraction between the poles for an infinitely short gap. This is, of course, the important force, because it determines the pull required to separate the poles.

When the air-gap is small, the flux-density (or field-intensity) in it is the same as in the iron, because the lines have no room to spread. The energy required to increase the gap from zero to δx is thus $Ea\delta x$, if E is the energy per unit volume and a the cross-section.

$$\text{But the energy required} = \text{Force} \times \text{distance moved,}$$
$$= \text{Force} \times \delta x.$$
$$\therefore \quad \text{Force} \times \delta x = Ea\delta x.$$
$$\therefore \quad \text{Force} = Ea$$
$$= \frac{B^2 a}{8\pi\mu}.$$

Let us imagine a concrete case.

Two electromagnets in a laboratory form two halves of an anchor-ring of diameter 20 cm. The cross-section of the iron is 8 cm.^{2}. Each electromagnet is wound with 100 turns of wire carrying 5 amperes. Find the energy stored and the force needed to separate the magnets, assuming that their poles are accurately flat and thus are in contact all over with the opposite poles. We will assume that the force can all be used to separate one pair of poles.

The permeability of the soft iron of the magnets is 2000.

We have seen from the last section that the self-inductance

$$L = \frac{2\mu a n^2}{r}$$
$$= \frac{2 \times 2000 \times 8 \times (200)^2}{10}$$
$$= 1{\cdot}28 \times 10^8 \text{ absolute units.}$$

The energy stored is thus

$$\tfrac{1}{2}Li^2 = \tfrac{1}{2} \, . \times 1{\cdot}28 \times 10^8 \times (0{\cdot}5)^2$$

(since 1 ampere = 0·1 absolute unit)

$$= 1{\cdot}6 \times 10^7 \text{ ergs.}$$

The energy per unit volume can be found in two ways.

First it is $$\frac{\text{total energy}}{\text{volume}} = \frac{1{\cdot}6 \times 10^7}{2\pi \times 10 \times 8}$$

$$= \frac{10^5}{\pi} \text{ ergs/c.c.}$$

Secondly it is $$\frac{\mu H^2}{8\pi} = \frac{2000 \times \left(\dfrac{2 \times 200 \times 0{\cdot}5}{10}\right)^2}{8\pi}$$

$$= \frac{10^5}{\pi} \text{ ergs/c.c.}$$

[The substitution for H is as before.

$$\left. H = \frac{4\pi ni}{\text{length of path}} = \frac{4\pi ni}{2\pi r} = \frac{2ni}{r}\right].$$

Now the force $= \text{Energy per unit volume} \times \text{cross-section}$

$$= \frac{B^2 a}{8\pi} = \frac{10^5}{\pi} \times 8 \times 2000 = 5{\cdot}09 \times 10^8 \text{ dynes}$$

Since in the indefinitely small air-gap between the poles $\mu = 1$ and $H =$ the value of B inside the iron.

Hysteresis

Hysteresis is the name for the lag of magnetization behind the magnetizing field. If a gradually increasing magnetic field H is applied to an unmagnetized piece of iron, its magnetization gradually increases until it reaches a maximum. If I for the iron is plotted against H, the magnetizing field, a curve like the single curve OA of Fig. 187 is obtained. The top of this curve is flat because H has been increased till the iron was saturated. If now H is gradually reduced, it is found that I reduces also, but not along the original curve.

It is clear that this must happen, for we know very well that if we magnetize a specimen of iron, and remove the magnetizing field, the residual magnetism is left. The I–H curve thus

N

returns along the line AB, and OB represents the residual magnetism.

If H is now made negative and increased in numerical value, I is at last reduced to zero at the point C in the graph. OC then represents the "coercive force," or reverse field required to destroy the residual magnetism.

H may now be made still more negative, until the iron is magnetically saturated in the opposite direction. This part of the curve is shown as CD.

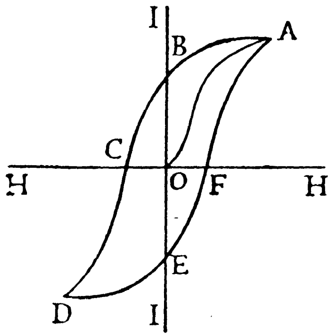


FIG. 187.

If H is now taken back from its negative saturation value to its original positive saturation value, a similar curve DEFA will be traced. The whole graph ABCDEFA thus forms a closed loop, usually known as a Hysteresis loop.

When a magnetic field produces magnetism in iron, it does work. The work is done in separating the opposite poles which appear in the iron, or (alternatively) in producing a magnetic moment per unit volume, or intensity of magnetization, in the iron.

The second way of considering the work is the better one, because the first cannot be applied to an endless ring of iron, on which no poles can appear when it is magnetized along the axis of the iron forming the ring.

Suppose a field H produces an intensity I in a piece of hitherto unmagnetized iron.

Let us consider the work done per c.c. of the iron.

Consider a rod (Fig. 188) of length l and cross-section a, magnetized to intensity I by a field H.

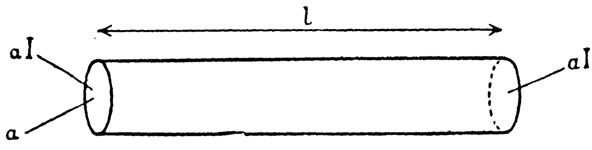


FIG. 188.

Poles aI and $-a$I appear at its ends.

Now let the magnetization increase to $(\mathrm{I} + d\mathrm{I})$.

The pole-strengths become $a(\mathrm{I} + d\mathrm{I})$ and $-a(\mathrm{I} + d\mathrm{I})$. Work has been done by H in transferring a quantity of pole $ad\mathrm{I}$ from one end to the other.

This work is $\mathrm{H}.a.l.d\mathrm{I}$.

Thus the total work done in changing the magnetization from I_1 to I_2 is

$$\int_{\mathrm{I}_1}^{\mathrm{I}_2} \mathrm{H}al\,d\mathrm{I}$$

which $$= al\int_{\mathrm{I}_1}^{\mathrm{I}_2} \mathrm{H}\,d\mathrm{I}.$$

Since al is the volume, work done per unit volume $= \int_{\mathrm{I}_1}^{\mathrm{I}_2} \mathrm{H}\,d\mathrm{I}$

Now let us consider our hysteresis curve.

The work done in taking the magnetization from one value of I to another is the area included between the bit of the curve concerned and the I-axis.

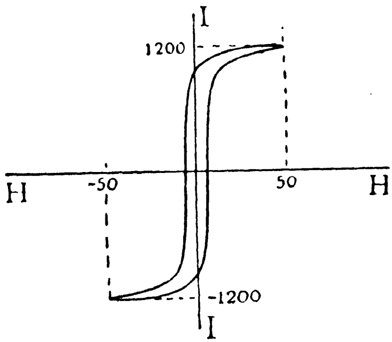


FIG. 189.

Thus the work done in going from positive saturation to negative saturation is different from that done in returning to positive saturation. The difference between these quantities of work is represented by the area inside the loop, and this quantity of work is irrecoverably lost during each cycle. It is lost only in the sense that it cannot be recovered or usefully employed. Its effect is to give heat to the iron and raise its temperature.

Since a field of H oersted does Hmx ergs of work in moving a pole m through a distance x, it follows that if H and I are measured in the same system of absolute units (always electro-

magnetic in practice, of course) the area enclosed in the hysteresis loop gives the energy lost per cycle per unit volume in ergs.

This loss varies from about 3000 ergs per c.c. per cycle for the most efficient silicon-iron in transformers to about 300,000 ergs per c.c. per cycle in hard steel.

It is obviously essential that in the cores of transformers, or in any iron which is given an alternating magnetization, the hysteresis loss must be made as small as possible. Otherwise two serious disadvantages operate: energy is lost, and the iron gets hot.

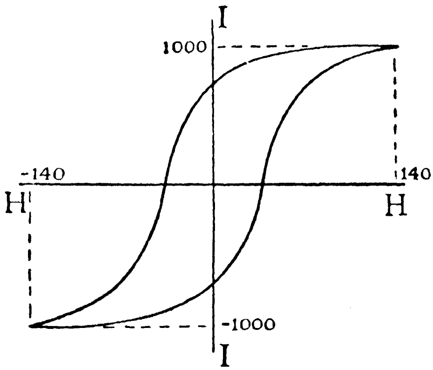


FIG. 190.

The energy is clearly lost owing to some cause which is very like friction. The magnetic elements of iron show some resistance to the action of a magnetic field either in magnetizing or demagnetizing the iron; and, like a frictional resistance, this opposes change in any direction.

Typical hysteresis curves for soft iron and hard steel are shown in Figs. 189 and 190.

Measurement of Hysteresis: Direct Method

This method is suitable for a specimen in the form of a cylindrical bar.

Two solenoids of wire are wound on cardboard cylinders just large enough for the bar to be slipped into them. One should be long enough to contain the bar and overlap it by an inch or so at each end. The other may be shorter. The number of turns to be wound depends upon the resources of the laboratory. It should be enough to produce an internal field of about 50 oersted without overheating the solenoid. For a maximum current of 1 ampere this gives about 40 turns per

cm. length, since $H = 4\pi ni$ when i is in e.m. units. About four layers of 22 gauge copper wire will probably do.

The apparatus is then arranged as in Fig. 191.

In this figure V is an accumulator, K a reversing switch, R a rheostat, A an ammeter, S the coil containing the specimen I, B the second coil, and M an ordinary tangent magnetometer.

The apparatus is arranged so that the axis of the coils is perpendicular to the horizontal component of the earth's field. The maximum current is switched on, and the specimen

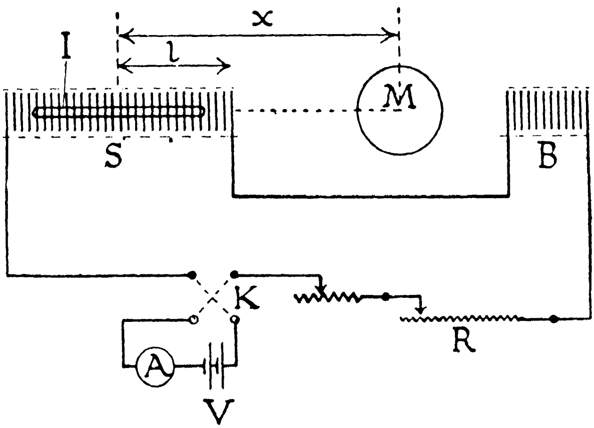


FIG. 191.

is put into its coil. The distance of the coil S is adjusted so that the magnetometer deflection is between 50° and 60°, in order that later in the experiment this deflection may not be found too large for accuracy. The specimen is now taken right away, and the position of the short coil B is adjusted so that its field at M is equal and opposite to that of S alone without the iron. When B is in the right place the magneto-meter should read zero for all values of current from zero to maximum. Any deflection obtained now when the iron is inside S will be due entirely to the iron and not at all to the solenoid S. B thus acts purely as a compensating coil.

The specimen should now be demagnetized.

The best way to demagnetize a piece of iron (without resort-

ing to such brutal methods as hammering or heating) is to apply to it a gradually diminishing alternating magnetic field.

If the specimen is put into the solenoid S, and the resistance R is gradually increased by small steps, the switch K being reversed after each increase, this will be done. Unfortunately a very small field produces considerable magnetization, and it is difficult and expensive to find a rheostat which will carry currents as big as 1 amp. and have a maximum resistance large enough.

It is quite a good plan to have two rheostats in series for R. One can have a maximum resistance of 20 ohms and carry 1 amp., and the other a maximum resistance of about 1000 ohms and carry up to $\frac{1}{10}$ amp. Then, if the second rheostat is short-circuited until all the resistance is employed in the first, it can be brought into action without being made to carry too much current.

When the specimen is demagnetized, it is placed in the solenoid S, and the current is given a number of values (say about 10) between 0 and 1 amp. For each value, θ is observed. The reversing switch is, of course, left untouched. The current is then brought back to zero by about the same number of steps, θ being observed. The switch K is reversed, and the current is increased again to the maximum, and then reduced back to zero, by steps as before. K is again reversed to its original setting, and the current is increased to its maximum. The deflection should now be the same, and in the same direction, as it was with the first maximum.

If $\tan\theta$ is plotted vertically against i, the current, horizontally, a curve of the general shape of the hysteresis curve should be obtained.

The value of H_0, the horizontal component of the earth's field, should now be determined. The simplest way to do this is by using a tangent galvanometer and a calibrated ammeter (which must, of course, be far enough from the galvanometer to cause no magnetic disturbance). Thus if a current i_0 is found to produce a deflection of $45°$ we have

$$i_0 = \frac{10rH_0}{2\pi n_0}$$

where r is the radius of the galvanometer coil and n_0 the number of turns.

Then $$H_0 = \frac{2\pi n_0 i_0}{10r}.$$

We must now find the relations between i and H, the magnetizing field inside S, and between $\tan\theta$ and I, the intensity of magnetization, respectively.

If no other influences acted, the field inside the solenoid S would be given by

$$H = \frac{4\pi n i}{10}$$

where n is the number of turns per cm. length of S and i is measured in amperes.

In our specimen, however, there is a demagnetizing effect due to the poles, since their field opposes the magnetizing field. This effect can be shown to be such that, if H is the effective field and H′ the calculated field from $H' = \dfrac{4\pi n i}{10}$, then

$$H = H'/\{1 + \alpha s\}$$

where s is the susceptibility and α a constant depending on the ratio of the specimen's length to its diameter.

Values of α are given in the following table, calculated by Clerk Maxwell—

$\frac{\text{Length}}{\text{Diameter}}$.	α.
50	0·0182
100	0·0054
200	0·0016
300	0·00075
400	0·00045
500	0·0003

Thus for a wire 400 times as long as it was broad, with a susceptibility of 200

$$H' = \frac{H}{1 + 200 \times 0{\cdot}00045}$$

$$= \frac{H}{1{\cdot}09}.$$

We thus see that this method is inaccurate for any but a long thin specimen.

Let us now consider I.

The magnetic moment is $\mathrm{I}v$, where v is the volume of the specimen in c.c.

If this produces a field F which causes the deflection θ, then

$$\mathrm{F} = \mathrm{H}_0 \tan \theta.$$

But by the end-on formula we have

$$\mathrm{F} = \frac{2\mathrm{I}vx}{(x^2 - l^2)^2}$$

if x is the distance from the magnetometer to the middle of the specimen of length $2l$.

Thus we have

$$\frac{2\mathrm{I}vx}{(x^2 - l^2)^2} = \mathrm{H}_0 \tan \theta.$$

$$\mathrm{I} = \frac{(x^2 - l^2)^2 \mathrm{H}_0 \tan \theta}{2vx}$$

But we have from the calculation for H

$$\mathrm{H} = \frac{4\pi ni}{10(1 + \alpha s)}.$$

Now energy lost per c.c. per cycle

$$= \int_{-\mathrm{I}}^{\mathrm{I}} \mathrm{H}d\mathrm{I}$$

$$= \left[\frac{\pi n(x^2 - l^2)^2 \mathrm{H}_0}{5vx(1 + \alpha s)}\right] \int_0 i\,d\,(\tan \theta)$$

since i and $\tan \theta$ are the only variables occurring in the equations for H and I.

∴ Hysteresis loss in ergs per cycle

$$= \frac{\pi n(x^2 - l^2)^2 \mathrm{H}_0}{5vx(1 + \alpha s)} \times (\text{Area of } i\text{-}\tan \theta \text{ loop})$$

where n = number of turns per cm. on solenoid S,
x = distance from magnetometer to mid-point of specimen,
$2l$ = length of specimen,
H_0 = horizontal component of earth's field,
v = volume of specimen,
α = Maxwell's constant,
s = approximate value of susceptibility of specimen.

The error due to the magnetizing effect of the vertical component of the earth's magnetic field has been neglected in this experiment.

Measurement of Hysteresis : Ballistic Method

A better method than the foregoing uses a ballistic galvanometer or fluxmeter to measure the induction B in a ring-shaped specimen having no free poles.

The circuit is connected up as in Fig. 192.

V is a battery, A an ammeter, K a reversing-switch, R a rheostat (or better two rheostats as in the case of the resistance R in the first method) which may be short-circuited by a

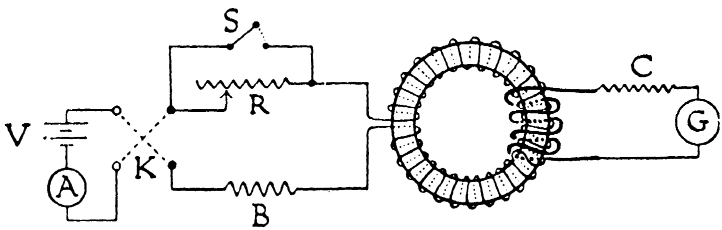


FIG. 192.

switch S. B is another resistance of such a value that when S is closed enough current flows in the circuit to produce magnet saturation. C is a resistance of whatever value is suitable for the ballistic galvanometer G.

The galvanometer must be calibrated for measurement of quantity and corrected for the damping due to the total resistance of the circuit, by the method of p. 391.

The left-hand circuit is connected to a coil uniformly wound round the whole ring, and producing a maximum magnetizing field of about 50 gauss. The CG circuit is in series with a number of turns of wire suitable for the sensitivity of the galvanometer.

The switch S is first closed, and saturation current is sent through the primary circuit. The reversing-switch K is thrown over, and thus the magnetization of the ring changes from positive saturation to negative saturation. The change of flux through the ring can be calculated from the observed

throw on the galvanometer, which should already have been calibrated for charge.

If n turns of wire are in the secondary, the change of line-linkages with the secondary circuit will be

$$2Ban$$

if a is the cross-section of the ring.

If a charge Q flows through the galvanometer we have in absolute e.m. units

$$Q = \frac{2Ban}{r}$$

where r is the total resistance of C, G, and the secondary coil.

Since we have calibrated the galvanometer, we know Q from the deflection.

Thus

$$B = \frac{Qr}{2an}.$$

Since $B = H + 4\pi I$, we have

$$I = \frac{1}{4\pi}\left(\frac{Qr}{2an} - H\right).$$

We have also, if the primary current is i absolute units, and there are N turns per cm. length of the ring,

$$H = 4\pi Ni.$$

Knowing Q and i, we can calculate I and H from these equations.

We thus have the two end points on the H–I curve.

Let us now imagine the final curve as already drawn (for convenience of explanation). We have at present the end points A and B (see Fig. 193).

A resistance of small value—say enough to reduce the current i by about one-third of its value—is given to the rheostat R.

If we have started at the point A, and used the reversing-switch once, we are now at the point B on the curve.

The switch S is opened, reducing the current, and a throw is observed which gives us the point C on the curve. The reversing-switch K is put over, and the throw gives us D on the curve. The switch S is closed and we should get back to A on

the curve. The switch S is opened and we get to E. The reversing-switch K is put over and we get to F. The switch S is closed and we get to B again. This process has now given us four points, C, D, E, F, on the curve besides the end points

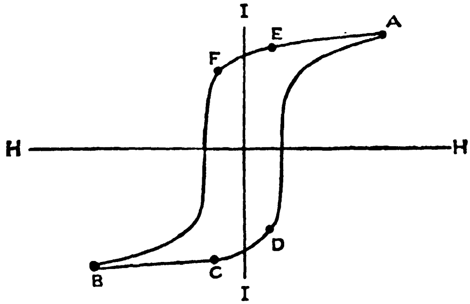


FIG. 193.

A and B; and it is repeated with various values of R until enough points are obtained to trace the whole curve.

A graph of I against H can then be plotted, and the hysteresis-loss can be calculated directly, since it is represented by the area included inside the H–I curve. It is of 20,000 ergs per c.c. per cycle for soft iron wire, but is much greater for steel wire.

CHAPTER VI

GALVANOMETERS

Types—Uses—Properties and Factor of Merit—Period—Damping—Zero-keeping Quality—Resistance—Optical System—Freedom from External Fields—Sensitivity—Factor of Merit—Moving Coil, Steady-current and Ballistic Galvanometers—Relation between Damping and Resistance—The Moving Magnet Galvanometer—Paschen Galvanometer—The Unipivot Galvanometer—General Principles—Constructional Details—Vibration Galvanometers—String, or Einthoven, Galvanometers—Einthoven String Galvanometer—Comparison of Sensitivities—The Grassot Fluxmeter.

Types

THE principal types of galvanometer are as follows :—

(*a*) Moving coil for steady current.
(*b*) Moving coil ballistic.
(*c*) Moving magnet.
(*d*) Unipivot.
(*e*) Vibration.
(*f*) String, or Einthoven.

The first two may be used either with a pointer or with a lamp and scale to magnify the deflection. In reflecting galvanometers a small mirror is fixed to the suspension. It reflects a bright source on to a scale. The galvanometer is usually calibrated with the scale at a fixed distance of 100 cm., but of course greater magnification of the deflection can be obtained by putting the scale further away and readjusting the focusing of the lamp.

A Moving Coil Galvanometer consists of a suspended coil which turns in a magnetic field when a current passes through it until the couple due to the field is equal to the restoring couple due to the twist of the suspension. It is used for measuring small steady currents.

A Ballistic Galvanometer is a moving coil galvanometer with

long period of swing and small damping coefficient. It is used for measuring the small charges moving round a circuit when a condenser is discharging or the flux is changing through a coil in a magnetic field.

A Moving Magnet Galvanometer consists of a small suspended magnet near a fixed coil carrying the current to be measured. The magnet turns until it is brought to rest by the restoring couple of the suspension. It has the advantage that its sensitivity can be varied by varying the number of turns of the fixed coil carrying the current.

A Moving Iron Galvanometer is similar in principle; but, as the iron is magnetized by the current in the coil, the deflection is always in the same direction. This type is of low sensitivity, and is used in cheap Ammeters and Voltmeters.

The Unipivot Galvanometer is a moving coil galvanometer in which the coil turns on one pivot instead of two. This system reduces friction and makes the levelling a much easier matter.

A Vibration Galvanometer is a moving coil galvanometer with a coil of very small moment of inertia which is capable of responding to electric oscillations of frequency of the order of 2000.

A String, or Einthoven, Galvanometer consists of a conducting fibre stretched in a strong magnetic field. If a momentary current passes the string is caused by the interaction of field and current to jump sideways. A point-image of the string, formed by a lens system on a moving photographic film, traces a line on the film, and shows sudden movements of the string as kinks in the line.

Uses

Among the chief uses of galvanometers are the following :—

(*a*) Measurement of small steady currents in the laboratory. [Moving Coil or Moving Magnet Reflecting Galvanometers of high resistance.]

(*b*) As the moving parts of ammeters or voltmeters or portable instruments. [Moving Coil or Moving Iron Pointer Galvanometers, or Unipivots.]

(*c*) Detecting uniformity of potential between two points as in Bridge methods of measuring resistance, and in Potentiometry. [Moving Coil or Moving Magnet Reflecting Galvanometers.]

(*d*) Detecting small potentials as in the earth inductor method of measuring the strength of the earth's magnetic field. [Moving Coil Ballistic Reflecting Galvanometers of low resistance. Strictly this method measures charge, but it may well be regarded as measuring potential, since a very small induced potential is causing the circulation of charge.]

(*e*) Measurement or comparison of small quantities of electricity, as in the measurement of magnetic flux or the comparison of the capacities of condensers. [Moving Coil Ballistic Reflecting Galvanometers.]

(*f*) Detection of very small intermittent rushes of current, and recording them, particularly in physiological work. [String Galvanometer.]

(*g*) Detection and measurement of small low-frequency alternating currents. [Vibration Galvanometer.]

Properties and Factor of Merit

The properties of a galvanometer to be considered when it is being chosen for a special purpose are as follows :—

(*a*) Undamped period.
(*b*) Damping.
(*c*) Zero-keeping qualities.
(*d*) Resistance.
(*e*) Optical system.
(*f*) Independence of external magnetic fields.
(*g*) Sensitivity.

Period

Generally speaking, for "zero," or "null" methods a galvanometer with a fairly short period (between 5 and 10 seconds) is useful, as although the sensitivity is not so high as that of one having a longer period, the spot returns quickly to zero, For ballistic work, a longer period is desirable. A short period galvanometer has many advantages, one being that observations may be rapidly taken and, if necessary, repeated for

verification, while the conditions of working are less liable to change during the test, owing to the short space of time occupied. Both the sensitivity and the period of a particular type of galvanometer depend on the controlling field. The sensitivity is proportional to the square of the period (undamped) when the period is varied by altering the control. The sensitivity may be increased by weakening the control, but only at the expense of quickness of working and stability of zero.

Damping

In the choice of a galvanometer, particular attention should be paid to the question of damping. To secure a maximum speed and ease of working, the galvanometer should be critically damped; that is, the moving system should on deflection reach its new position and come to rest there in the minimum time, and, as soon as the current is removed, return rapidly to zero without any reverse deflection. Some observers prefer the moving system to be slightly under-damped so as to cross the rest point once. If the galvanometer is unduly under or over-damped, the time taken for the system to come to rest is excessive; if it is over-damped, the movement is so sluggish as to lead to errors due to indefiniteness of zero. In suspended coil galvanometers, critical damping is effected by altering the resistance so as to adjust the magnitude of the induced currents. If the resistance of the apparatus with which the galvanometer is used is less than the external critical resistance (that is, the resistance external to the galvanometer requisite to ensure critical damping), additional resistance is added in series with the galvanometer. If the resistance of the circuit is too great, critical damping is usually obtained by connecting a suitable resistance in parallel with the galvanometer coil. In all cases a critically damped galvanometer is so much easier to manipulate that it is frequently advisable to sacrifice the amount of sensitivity necessary to give it this quality. In moving coil instruments the damping is sometimes controlled by varying the magnetic field. Other factors being constant, electro-magnetic damping is proportional to the square of the magnetic field.

The damping discussed here is, of course, electro-magnetic

only. Frictional damping also occurs. Both kinds are discussed in the section on the equations of moving coil galvanometers.

Zero-keeping Quality

This depends largely on the control used; generally speaking, the weaker the control the more unstable the zero. In moving coil galvanometers the zero-keeping quality also depends on the absence of magnetic impurities in the coil. In zero methods, where the galvanometer is merely used as a detector, variations of zero are not of primary importance, but if the galvanometer is to be used quantitatively a stable zero is essential. The control may, therefore, be made weaker for zero work, with a consequent gain in sensitivity, provided that the period does not become too long.

Resistance

In a galvanometer of a particular type, where the winding space is already determined, the sensitivity corresponding to a given current is approximately proportional to the two-fifths power of its coil resistance, and inversely proportional to the three-fifths power of its coil resistance for a given potential difference applied to the terminals, neglecting the effect of the resistance of the suspension in the case of a moving coil instrument. Consequently the higher the resistance of a galvanometer, the more sensitive it is for a given current, and the lower the resistance the more sensitive it is for a given potential difference at its terminals. Generally speaking, when great current sensitivity is required, as when measuring high insulation, a high resistance instrument should be used, while when it is desired to measure very small electromotive forces, as, for instance, in thermo-electric measurements and in low resistance bridges, a low resistance galvanometer will be found most sensitive. The best sensitivity in a bridge or potentiometer is obtained when the resistance of the galvanometer is comparable with the other resistances, although to secure critical damping and quickness of reading its value may have to be made lower.

In a moving coil instrument with low resistance coil, the resistance of the suspension strip may become important, and in this case there is no advantage in reducing the resistance of

the coil below a certain point. This limiting value is 20 ohms in most cases. There may, however, be special cases in which very strong controls, having a resistance of about one or two ohms, are used, and then a galvanometer of from 7 to 10 ohms resistance may be an advantage. It should be remembered that for a given constant electromotive force in circuit, the resistance of the coil should, if possible, equal all the resistance of the external circuit, including the suspension. When the moving coil is of low resistance, the sensitivity of the instrument to small electromotive forces is limited largely by the resistance of the suspension. By choosing a material rather for its low resistance than for its freedom from elastic fatigue (for example, by using silver-gilt strip instead of phosphor bronze), higher sensitivity may be obtained at the expense of stability of zero.

Optical System

The precision with which measurements may be made depends largely on the magnification and definition of the optical system. The mirrors and lenses must be optically correct for the working distance in order to give a well-defined spot, and they must also be of large area to reflect sufficient light for easy reading. For most work, a convenient scale distance is one metre, but, by increasing this distance, the linear deflections will be proportionally magnified. From the point of view of convenience and absence of fatigue in reading, the use of a transparent scale and illuminated spot is to be recommended. The focusing of the spot can be effected by the use of a concave mirror on the moving system, or by using a plane mirror in conjunction with interchangeable lenses; the latter method is preferable where the scale distance may have to be changed. A reading telescope used with an opaque scale and a plane mirror on the galvanometer has the advantage of requiring no separate source of illumination; it affords the possibility of considerable magnification of the readings at a comparatively short working distance.

Freedom from External Fields

If a galvanometer is affected to any considerable extent by external fields, its readings will be unreliable. Instruments

which are liable to be so affected should be protected by suitable shields. External temperature variations in general do not affect a galvanometer; special precautions are taken to reduce their effect to a minimum in cases where it would interfere with the accuracy and stability of the instrument.

Sensitivity

The sensitivity of a galvanometer has been expressed in a variety of ways. The current sensitivity has been defined (1) as the reciprocal of the current required to produce a deflection of one millimetre on a scale placed one metre from the galvanometer mirror, and (2) as the deflection on the scale produced by unit current. The most modern definition of current sensitivity is, however, the deflection in millimetres produced on a scale one metre away by a current of one micro-ampere. The voltage sensitivity is usually defined as the deflection obtained on a scale one metre away when the voltage applied to the instrument is one micro-volt. The voltage sensitivity is, therefore, the current sensitivity divided by the resistance of the galvanometer.

Factor of Merit

Ayrton and Mather have pointed out that the sensitivity of a galvanometer for a given direct current is proportional to

(*a*) the square of its period (undamped) when the period is altered by altering the control.

(*b*) to the two-fifths power of its coil resistance, the coil volume being constant.

When comparing galvanometers it is therefore advisable to reduce the sensitivities to the values they would have if all the galvanometers had unit period and resistance. For any galvanometer, this value, called the Factor of Merit, is :—

$$\frac{100 \times D}{T^2(R)^{\frac{2}{5}}} \text{ or } \frac{100\, D_1 R^{\frac{3}{5}}}{T^2}$$

where T = Periodic time (undamped) in seconds,

R = Galvanometer resistance in ohms,

D = Deflection in millimetres per micro-ampere at a scale distance of 1 metre,

D_1 = Deflection in millimetres per micro-volt at a scale distance of 1 metre.

Assuming negligible damping, the sensitivity of a galvanometer per micro-coulomb can be calculated from the micro-ampere sensitivity from the formula

$$D_2 = \frac{2\pi D}{T}$$

or, with small damping with logarithmic decrement λ,

$$D = \frac{D_2 T}{2\pi}\left\{1 + \frac{\lambda}{2}\right\}$$

where D = Deflection in millimetres per micro-ampere at a scale distance of 1 metre,
T = Periodic time of swing (seconds),
D_2 = Deflection in millimetres per micro-coulomb.

The sensitivity and also the period required must be known before a suitable galvanometer can be selected.

Moving Coil, Steady-current and Ballistic Galvanometers

The correct understanding and use of these galvanometers depends considerably upon having some understanding of their mathematical theory, which is now given.

Let i be the current at time t.
Let H be the magnetic field in which the coil hangs.
Let A be the area of the coil.
Let n be the number of turns of wire in the coil.
Let θ be the angle through which the coil has turned at time t.
Let μ be the restoring couple of the suspension for a twist of 1 radian, so that $\mu\theta$ is the restoring couple for twist θ.
Let i_0 be the steady current which produces a deflection θ_0.
Let T be the time of a complete oscillation of the suspended coil.
Let I be the moment of inertia of the moving coil.
Let d be the decrement.
Let λ be the logarithmic decrement.
Let Q be the total charge passing which produces a deflection θ at the end of the first swing, θ_2 at the end of the second to the opposite side, θ_3 at the end of the third to the original side, and so on.

It is assumed that the whole of the charge has passed before the coil has had time to move an appreciable distance, so that

the maximum angular momentum is attained when θ is still appreciably zero.

Let electromagnetic couple when the current is i be

$$\mathrm{HA}ni.$$

We have therefore by the momentum equation

$$\text{couple} = \text{rate of change of angular momentum}$$
$$\int \text{couple}\, dt = \text{total angular momentum.}$$
$$\mathrm{I}\frac{d^2\theta}{dt^2} = \mathrm{HA}ni$$

Integrating

$$\mathrm{I}\frac{d\theta}{dt} = \mathrm{HA}n\int idt, \text{ which is the whole impulse.}$$
$$\therefore \quad \mathrm{I}\frac{d\theta}{dt} = \mathrm{HA}n\mathrm{Q} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Since $\int idt$ taken as long as the current flows is Q.

The kinetic energy given to the coil at the start is

$$\tfrac{1}{2}\mathrm{I}\left(\frac{d\theta}{dt}\right)^2$$

which from our equation (1) already obtained is

$$\tfrac{1}{2}\frac{(\mathrm{HA}n\mathrm{Q})^2}{\mathrm{I}}.$$

At the same time the work done in giving the coil a twist of θ against the torsion of the suspension is

$$\int_0^{\theta_1} \mu\theta d\theta = \tfrac{1}{2}\mu\theta_1^2.$$

Since the kinetic energy is destroyed in doing this work,

$$\tfrac{1}{2}\frac{(\mathrm{HA}n\mathrm{Q})^2}{\mathrm{I}} = \tfrac{1}{2}\mu\theta_1^2$$

so

$$\mathrm{Q}^2 = \frac{\mu\mathrm{I}}{\mathrm{H}^2\mathrm{A}^2n^2}\cdot\theta_1^2 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

At the same time we have that the time of swing T of the coil of moment of inertia I with a torsional restoring couple $\mu\theta$ is given by

$$\mathrm{T} = 2\pi\sqrt{\frac{\mathrm{I}}{\mu}}$$

so

$$\mathrm{I} = \frac{\mu\mathrm{T}^2}{4\pi^2}.$$

Substituting for I in equation (2)

$$Q^2 = \frac{\mu^2 T^2}{4\pi^2 H^2 A^2 n^2} \theta_1^{\,2}$$

or
$$Q = \frac{\mu T}{2\pi H A n} \theta_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

At the same time since the electromagnetic couple for current i_0 is $HAni_0$, and the torsional couple for θ_0 is $\mu\theta_0$

$$HAni_0 = \mu\theta_0.$$

so
$$\frac{\mu}{HAn} = \frac{i_0}{\theta_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

(These two equations apply of course to steady deflections only.) So substituting (4) in (3)

$$Q = \left(\frac{i_0 T}{2\pi\theta_0}\right)\theta_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

In this equation i_0, θ_0, T and θ_1 are all determined directly; so Q can be found.

This equation, however, does not allow for the frictional damping.

If the coil is allowed to swing freely, its swings get progressively shorter and shorter owing to frictional resistance. It is found experimentally that each swing bears the same ratio to the one before; *i.e.* that

$$\frac{\theta_1}{\theta_2} = \frac{\theta_2}{\theta_3} = \frac{\theta_3}{\theta_4} = \ldots$$

This ratio is called the decrement, *d*.

λ, the logarithmic decrement, is the log of *d* to base *e*. That is,

$$\frac{\theta_1}{\theta_2} = \frac{\theta_2}{\theta_3} = \frac{\theta_3}{\theta_4} = \ldots = e^{\lambda}.$$

The swing we are calculating equation (5) for was an undamped swing. We should therefore allow for damping by finding what this swing would have been if the frictional effect had been cut out.

Half a period passes between θ_1 and θ_2. So clearly quarter

of a period passed while the coil swung out from zero to θ_1. Let us call the undamped throw θ_{max} —

So since
$$\theta_1 = \theta_2 e^{\lambda}$$
$$\theta_{max} = \theta_1 e^{\frac{\lambda}{2}}$$
$$= \theta_1\left(1 + \frac{\lambda}{2} + \frac{\lambda^2}{4\underline{|2}} + \frac{\lambda^3}{8\underline{|3}} + \cdot \quad \right).$$

As λ is small we may take without appreciable error—

$$\theta_{max} = \theta_1\left(1 + \frac{\lambda}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The accurate form of equation (5) is thus found by substituting θ_{max} for θ_1. It then becomes

$$Q = \left(1 + \frac{\lambda}{2}\right)\left(\frac{i_0 T}{2\pi\theta_0}\right)\theta_1. \quad . \quad . \quad . \quad . \quad (7)$$

λ can easily be calculated from observations of θ_1, θ_2, θ_3, etc.

Relation between Damping and Resistance

If the circuit of a galvanometer coil swinging freely be closed instead of left open, it will according to Lenz's Law be damped electromagnetically as well as by friction. The current induced in the coils by moving across the field will experience a force due to the field opposing the motion. Electromagnetic damping is used to bring the coil quickly to rest after a reading.

The motion of the coil is then affected by

(1) The couple due to the twist of the suspension. This acts to decrease θ, and is equal to $\mu\theta$.

(2) The friction couple, supposed proportional to $\frac{d\theta}{dt}$, the angular velocity. Call it $p\frac{d\theta}{dt}$.

(3) The electromagnetic couple, directly proportional to $\frac{d\theta}{dt}$, and inversely proportional to R, the *total* resistance in the circuit. Call it $\frac{m}{R}\frac{d\theta}{dt}$.

The equation of motion is therefore

$$\mathrm{I}\frac{d^2\theta}{dt^2} = -\left(\frac{m}{\mathrm{R}} + p\right)\frac{d\theta}{dt} - \mu\theta.$$

$$\therefore \quad \mathrm{I}\frac{d^2\theta}{dt^2} + \left(\frac{m}{\mathrm{R}} + p\right)\frac{d\theta}{dt} + \mu\theta = 0.$$

Let $$\frac{\frac{m}{\mathrm{R}} + p}{\mathrm{I}} = 2b, \text{ and let } \frac{\mu}{\mathrm{I}} = k^2.$$

The equation becomes

$$\frac{d^2\theta}{dt^2} + 2b\frac{d\theta}{dt} + k^2\theta = 0.$$

(This extremely important equation occurs frequently in problems of mechanical, electrical, or acoustic oscillations.)

The ordinary solution of this is

$$\theta = \mathrm{A}e^{-bt}\cos\sqrt{k^2 - b^2}t.$$

When $t = o$, $\theta = \theta_1$ (the most convenient time to start counting).

Then $$\theta = \theta_1 e^{-bt}\cos\sqrt{k^2 - b^2}t \quad . \quad . \quad . \quad . \quad . \quad (8)$$

After this the successive maxima θ_2, θ_3, θ_4, etc., occur very nearly when $[\sqrt{k^2 - b^2}t]$ is π, 2π, 3π, etc., since these values make $\cos\sqrt{k^2 - b^2}t$ equal to -1, 1, -1, etc.

They are not *quite* at these points, but are near enough for practical purposes.

Substituting these values of t in (8) we get

$$\theta_2 = \theta_1 e^{\frac{-b\pi}{\sqrt{k^2 - b^2}}}$$

$$\theta_3 = \theta_1 e^{\frac{-2b\pi}{\sqrt{k^2 - b^2}}}$$

$$\theta_4 = \theta_1 e^{\frac{-3b\pi}{\sqrt{k^2 - b^2}}}$$

and so on.

So the logarithmic decrement d is given by

$$d = \frac{\theta_1}{\theta_2} = \frac{\theta_2}{\theta_3} = \frac{\theta_3}{\theta_4} = e^{\frac{b\pi}{\sqrt{k^2 - b^2}}}.$$

So we have

$$\lambda = \frac{b\pi}{\sqrt{k^2 - b^2}}.$$

Actually this solution does not hold when $k = b$. We may, however, consider the case when k is nearly equal to b.

As $k \to b$, $\lambda \to \infty$, and the time of coming to rest approaches a minimum. Hence this is the condition wanted in all galvanometers measuring steady currents.

But a ballistic galvanometer is one in which both damping couples are small compared with the torsional couple (for it must swing freely) as well as one in which I is very small (for it must start easily).

Its time of swing must be great, so that the whole impulse is given before it starts. So we may neglect b^2 compared with k^2, and write k for $\sqrt{k^2 - b^2}$.

Thus
$$\lambda = \frac{b\pi}{k}$$
$$= \left(\frac{\frac{m}{R} + p}{2\sqrt{\mu}}\right)\pi.$$

As R is the only quantity we can measure accurately we may as well write this equation as

$$\lambda = C\left(\frac{1}{R} + p\right)$$

where C and p are simply regarded as constants. This equation should be able to be directly verified, specially with a low resistance galvanometer. It could even be used to determine the galvanometer resistance by varying R by giving the external resistance a series of values from zero to about double that of the galvanometer, and plotting $\frac{1}{R}$ against λ.

A good example of a dual-purpose galvanometer is the

Cambridge Instrument Company's High Sensitivity Galvanometer, shown in Fig. 194, which may be used either for steady currents or ballistically.

It may be fitted with a coil having either a high or a low resistance, the decision as to which coil is required depending upon the purpose for which the galvanometer is to be used. When fitted with a high resistance coil, the instrument is particularly suitable for the measurement of high insulation resistances; when fitted with a coil of low resistance, it gives a high volt sensitivity such as is required for measuring minute differences of temperature with thermo-couples. By connecting in circuit a suitable form of external thermal accessory such as an independent junction or a vacuo-junction, the galvanometer can be employed for the measurement of alternating currents down to a few microamperes, this combination being almost independent of frequency.

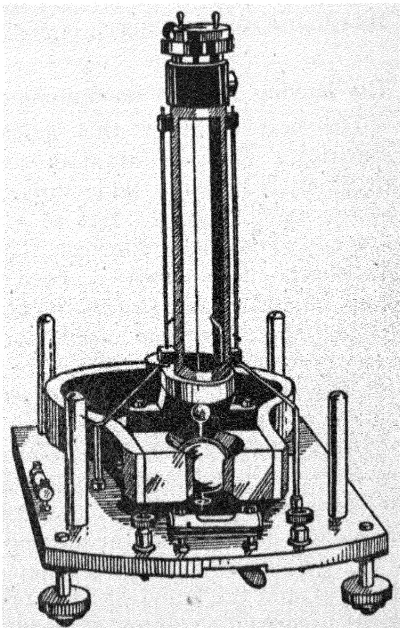


FIG. 194.

The moving coil is suspended by a single flat strip of phosphor bronze, 22 centimetres long, the connections to the phosphor-bronze strip being made through two silver leading-in spirals. Electromagnetic damping may be applied by shunting the coil with a suitable resistance, but when the instrument is to be employed for ballistic work, or, having a low internal resistance, is to be used on a very low resistance shunt, this electromagnetic damping is not required.

The galvanometer is usually fitted with a concave mirror

10 millimetres in diameter, which gives a sharp image on the scale at a distance of one metre. The standard instrument is fitted with a coil having an approximate resistance of 15 ohms, but instruments with coils having approximate resistances of 20, 300 or 2,000 ohms respectively can also be supplied. The external damping resistance is approximately twenty times the coil resistance of the instrument; for example, in the case of the 15-ohm galvanometer, an external resistance of 300 ohms is required in order to obtain critical damping.

The Moving Magnet Galvanometer

The best form of this galvanometer gives very high sensitivity if it is wanted, and the sensitivity can be easily varied over a wide range. When the sensitivity is of the same order as that of a moving coil galvanometer, the period is much shorter. The suspension can be made of quartz fibre since it need not be conducting. This kind of suspension suffers much less from elastic fatigue. Since the suspension need not carry current, the total galvanometer resistance can be made much smaller than for the moving coil, and the sensitivity to small E.M.F.s can be greatly increased.

In the Paschen instrument the sensitivity can be varied without touching the moving system by arranging the coils in series, series-parallel, or parallel. It is, however, more delicate to handle than the moving coil type.

The Paschen moving magnet galvanometer is shown in Fig. 195, and the following extract gives the Cambridge Instrument Company's account of their instrument.

"Paschen Galvanometer

The galvanometer illustrated in Fig. 195 is the result of investigations carried out by Paschen, Mendenhall and Waidner, and Abbott, and incorporates recent modifications due to Prof. A. V. Hill and Mr. A. C. Downing which have increased the high sensitivity, astaticism and factor of merit characteristic of the original design.

"The moving system consists of two groups of three cobalt

steel magnets mounted on a glass stem which also carries a small mirror. The magnets are so made as to ensure uniformity of size and physical condition. The magnet system, which is suspended by a quartz fibre about 0·0001 mm. in diameter and 20 centimetres in length, may be readily replaced. The four coils are wound on elliptical cores; they can be connected readily in series, series-parallel, or parallel, the approximate resistances in the standard instrument, when so connected, being 12 ohms, 3 ohms and 0·75 ohm respectively. They may be easily removed and replaced. When the galvano-

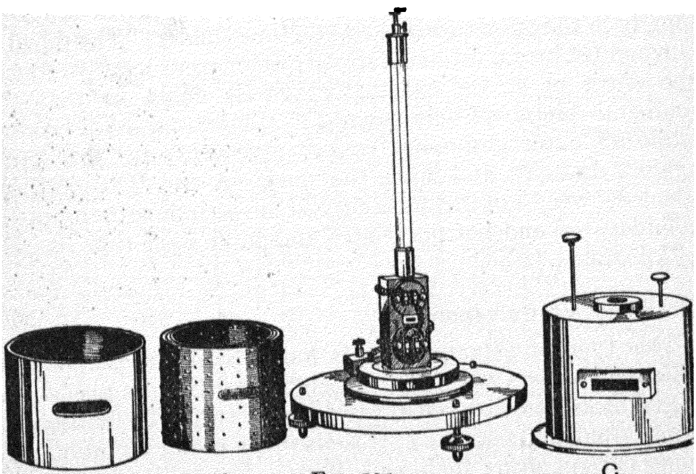


FIG. 195.

meter is to be used with a high resistance external circuit, such as high resistance thermopiles, a set of four coils, each having a resistance of about 750 ohms, can be supplied. These coils give a range of resistance on the galvanometer from about 180 to 3000 ohms. The standard optical system consists of a plane optically-worked mirror 1·5 millimetres square, and a lens of 1 metre focal length. The body of the instrument is machined from a single casting on which are rigidly mounted the coils and the glass suspension tube. A large diameter base simplifies the levelling of the galvanometer.

"The resistance, sensitivity and period may be readily varied by interchanging the coils and by altering the magnetic

control. The galvanometer is particularly suitable for use with thermopiles, for taking measurements in calorimetry and radiometry, and it should be employed for measurements requiring the highest order of sensitivity, astaticism and period control.

"The moving system is protected from external magnetic fields by a cylindrical shield made of high permeability alloy known as 'mumetal' which is exceptionally light in relation to its shielding powers, and is so designed as to secure permanent efficiency. The cylinder (A) is closed at either end by mumetal plates; the bottom plates are so positioned that no stress is transmitted to them, and the top plates are held loosely in the outer case and rest on the cylinder. The shield is surrounded by a mild steel tube (B) which considerably reduces the effect of strong external magnetic fields, so that the mumetal shield will only require to eliminate a small residue, while an outer gun-metal case C gives effective protection against draughts and holds the shields A and B in position. The total weight of the shielding system, including the mumetal cylinder, top and bottom plates and mild steel tube, is only $3\frac{1}{2}$ kilograms."

The Unipivot Galvanometer

This type of galvanometer is most useful for purposes in which robustness without very high sensitivity is needed, or where measurements must be made, in spite of vibration, in ships or moving vehicles. It is also very useful for laboratory work not requiring high sensitivity. The following detailed description of a unipivot galvanometer is by the Cambridge Instrument Company.

"General Principles

"In Unipivot instruments for use with direct current measurements a circular coil swings concentrically round a spherical iron core mounted between the poles of a permanent magnet. The coil is fitted with a vertical spindle (Fig. 196) provided with a fine steel pivot which rests in a jewel bearing at the centre of the core. The current is led in through the cylindrical spring H (Fig. 197) and out through the flexible ligament L. The pointer P and the counterweights are attached to the top of the spindle, and by accurate balancing, the centre

of gravity of the moving system is made to coincide with the point of the pivot. The absence of any appreciable friction renders the instrument extremely sensitive, while, owing to the mechanical strength of the moving parts, it is of robust construction and is not easily injured. On account of the damping action of the pointer, the oscillations of the moving coil do not affect the readings when the instrument is used in positions subject to considerable vibration, and it can therefore be used without difficulty near running machinery. As it is unneces-

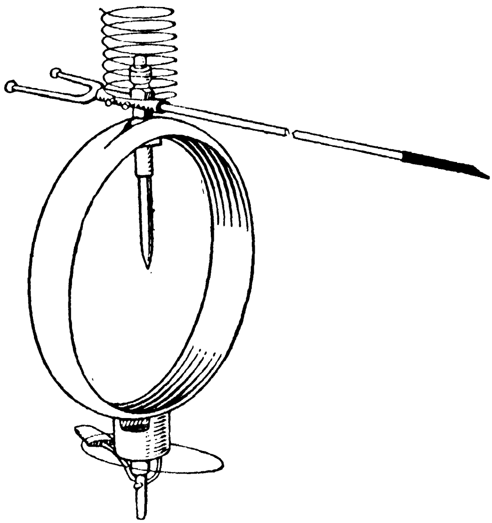


FIG. 196.

sary to level the instrument, it is suitable for use on board ship or in rapidly moving vehicles.

"**Constructional Details**

"*Moving System.* The spherical core S (see Fig. 197) is drilled and tapped and fitted with a hollow screw; in this the jewel is mounted on a spring of such strength as to yield, and allow the coil to stop against the core, before any injurious pressure can be applied to the pivot. The magnet is either of circular form, as shown in Fig. 198, or a powerful horseshoe magnet is used which is fitted with soft iron pole-pieces, bored from the

solid. The flexible phosphor-bronze ligament L below the coil exerts no appreciable control on the movement and serves merely to lead the current from the coil to the negative terminal of the instrument. The movement of the pointer is controlled by the spring H, the upper end of which is held in a geometric clamp in the zero-adjusting device Z. This arrangement enables the effective length of the spring to be easily altered, permitting accurate adjustment of the movement

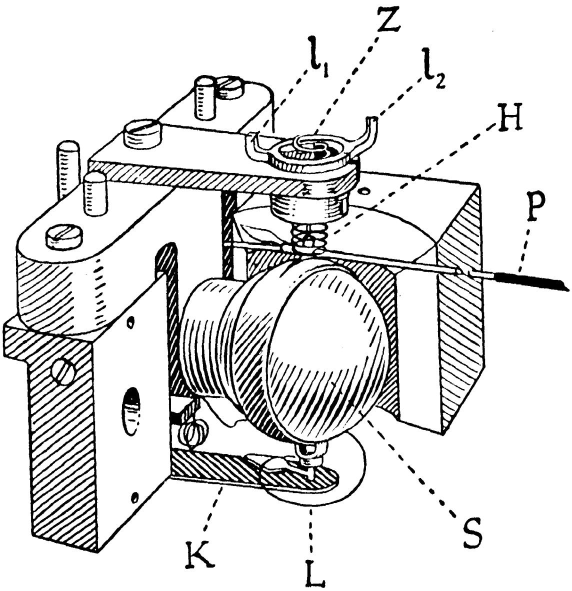


FIG. 197.

to the required sensitivity without deforming the spring. After this adjustment has been made, the connection is soldered. A simple lifting device K enables the entire weight of the coil to be removed from the pivot when the galvanometer is not in use. This consists of a flat spring cupped at a point below the vertical axis of the coil and pivoted on a fulcrum near its centre. On depressing a clamp on the face of the instrument, the cup engages the lower spindle of the coil, lifting it off the jewel, and a slight rotation of the clamp maintains the coil in this position.

"*Damping.* The moving coil is wound on a metal former which damps the movement of the coil without making it sluggish. This leaves the instrument slightly underdamped on open circuit, which is desirable, since working conditions introduce additional damping. In this way the quickest working speed and dead-beat action are attained. The pointer swings a little past its final position before coming to rest; this is a check on the absence of friction and increases the speed and certainty of reading. For ballistic work, the instruments are supplied undamped.

"*Zero Adjustment.* The zero-adjusting device Z (Fig. 197) is of simple form. It consists of a collet which rotates on a

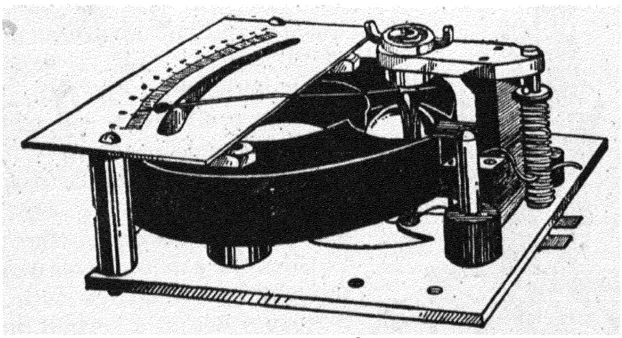


FIG. 198.

tube carried by a projection from the cock-piece on which the core is mounted. The geometric clamp which holds the control spring H forms part of this collet. Two projecting ɪgs l_1, l_2 (Fig. 197) engage with a fitting in the cover plate of ·e instrument, and the zero of the instrument is adjusted rotating a stud which is attached to this fitting.

'*Sensitivity.* The sensitivity of Unipivot galvanometers · be specified as normal, high, or low, according to the ·ol spring fitted. In instruments of normal sensitivity, will be found to meet ordinary commercial requirements, ·iod varies between three and five seconds, according ›attern supplied. In instruments of low sensitivity a control spring is used with heavy damping and the ; reduced to about two seconds. The highest sensi-

tivity is obtained by fitting a spring suspension to reduce pivot friction; the period is then about nine seconds."

Vibration Galvanometers

These galvanometers have very small moving coils which can be adjusted to oscillate with any desired natural frequency over ranges between about 10 and 2000 oscillations per second. They can thus be used to measure alternating currents with which they have been brought into resonance. Their resonance curve is in general very sharp, so that they respond only to one frequency, and take almost no notice of all others. Such a galvanometer is shown in Fig. 199. This has a narrow moving coil, of very small moment of inertia, which is held by a double bifilar suspension. By varying the effective length and tension of the suspension, the galvanometer may be brought into resonance with any frequency within a certain range. The optical system consists of a plane mirror and a convex lens of 100 cm. focal length. Though the frequency range can be extended from 10 to 2000 by changing coils it can only vary over part of this range for any one coil. In tl galvanometer shown, standard coils cover ranges from 10 100, 20 to 200, 30 to 300, 60 to 600, or 200 to 1200 cycles second.

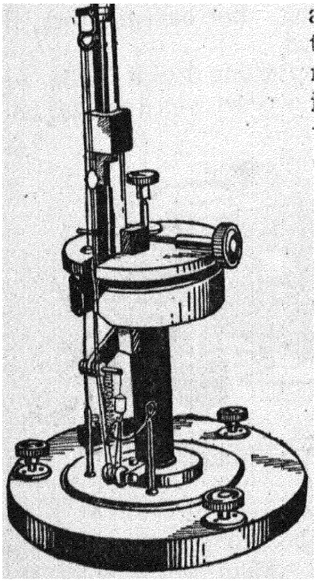


FIG. 199.

The coils will of course go outside these ranges. range of the last coil, for example, may be extended tc cycles per second.

String, or Einthoven, Galvanometers

The principle of these galvanometers is completely from that of all others. They use directly the fac

conductor carrying a current in a magnetic field experiences a force perpendicular to itself and the field. A thin wire is stretched across a magnetic field. If a momentary pulse of current runs through the wire it kicks sideways in response. A point-image of the wire is thrown, by means of a system of lenses, on a moving film, where it traces a white line on a black background. This line shows a kink whenever a pulse of current passes. The galvanometer is described in detail in the following extract from the Cambridge Instrument Company's catalogue.

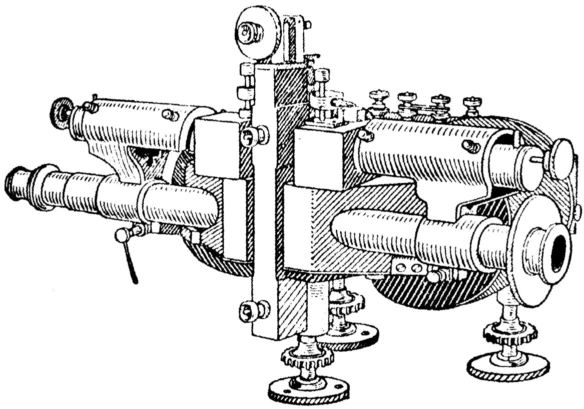


FIG. 200.

"**Einthoven String Galvanometer**

"This galvanometer (Fig. 200), originally designed by the late Professor Einthoven, is extremely sensitive, dead-beat, has a short period, and possesses practically no self-inductance or capacitance. In conjunction with a recording camera, alternating currents of frequencies up to 200 per second may be accurately recorded. The instrument is also valuable for physiological investigations, in particular for recording the changes of electrical potential which precede the contractions of the human heart.*

"The galvanometer is of the moving-coil type, the coil,

* *The Mechanism of the Heart Beat*, by Sir Thomas Lewis, M.D., F.R.S.

mirror and suspension being replaced by a single fine wire, or conductive fibre, which is stretched in a narrow air-gap between the poles of a powerful electromagnet. When a current passes through the fibre the latter is deflected at right angles to the magnetic field; the deflection is observed through a hole bored in the pole-pieces, either by means of a microscope or by projecting an image of the fibre on to a screen or photographic film.

"The fibre is generally of glass covered with a thin metallic coating, and is approximately 65 millimetres long and from 0·002 to 0·005 millimetre in diameter. Fibres of these dimensions vary in resistance from 1400 to 6000 ohms, depending on the thickness of the coating; the resistance of a particular fibre, however, remains remarkably constant. Alternatively, the standard fibre may be replaced by a fine metal wire. The

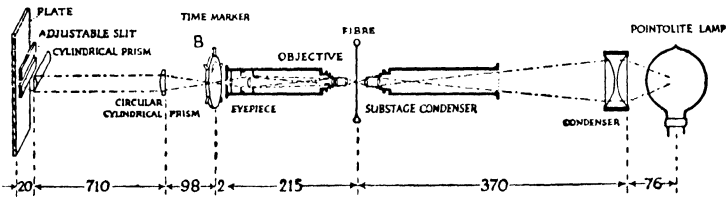


FIG. 201.

fibre is completely enclosed in a removable case, fitted with mica windows, through which the movement is observed; a fine motion screw enables the fibre tension to be adjusted and the sensitivity and period thereby controlled, a stop being fitted to prevent overtightening. The fibre case can be easily withdrawn and quickly replaced by a spare fibre case with fibre already mounted. A double fibre case can alternatively be fitted, in which two fibres are mounted side by side, enabling simultaneous records of two separate electrical phenomena to be obtained on one photographic film or plate.

"The electromagnet is so shaped as to concentrate the field in the air gap in which the fibre moves, and small changes in the field current have little effect on the intensity of the saturated magnetic field around the fibre. The field coils which energise the magnet are wound in two sections which can be connected either in series or in parallel, thereby making

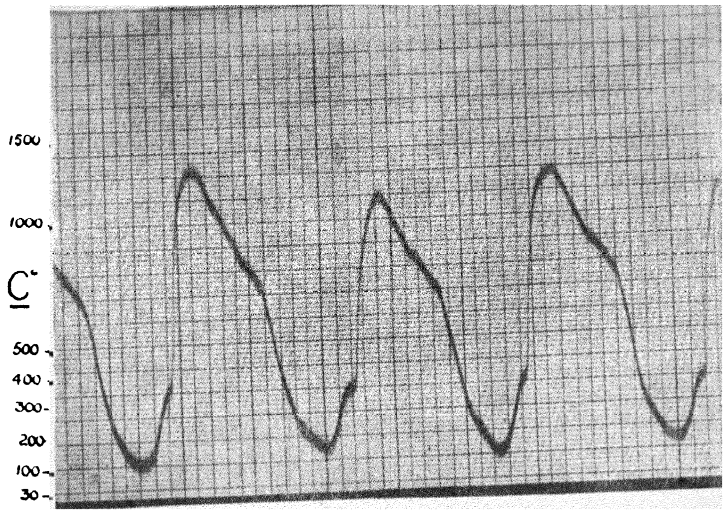


FIG. 202.

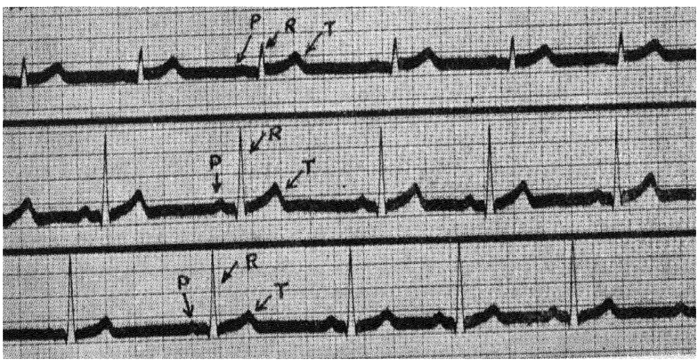


FIG. 203.

each instrument suitable for two voltages. The standard windings are for 110 or 220 volts and for 12 or 24 volts respectively. In either case, the energy consumed is about 40 watts.

"The optical system is supported from the electro-magnet by geometric fittings, which permit of it being accurately aligned and focused. The arrangement of the optical work

depends on whether the deflections of the fibre are to be (1) observed by means of a microscope; (2) projected on a screen; or (3) recorded photographically. The latter method is usually adopted, the optical work being arranged as in Fig. 201. Light from a Pointolite lamp is concentrated on the fibre, the image of which is focused on to a sensitised plate. Unless otherwise specified, an 11-millimetre apocromat objective and a two-third inch eyepiece are employed.

"Owing to the considerable magnetic and air-damping, the moving system does not overshoot, while the zero remains constant indefinitely, provided that excessive slackening of the fibre is avoided. The magnetic damping may be varied by altering the current in the magnet field coils or the resistance in circuit with the fibre."

Typical records are shown in Figs. 202 and 203.

Comparison of Sensitivities

The table shown below gives some information about the

	Coil Resistance (ohms).	Period of Swing (seconds).	Deflection in mms. at 1 metre.			Factor of Merit.
			Per microamp.	Per microvolt.	Per micro-coulomb.	
Moving Coil (High Sensitivity)	15	15	300	20	126	45
	15	22	600	40	171	40
	300	22	4,000	13	1,142	85
	2,800	22	12,000	4·3	3,430	100
Paschen Moving Magnet	12	6	14,600	1,220		15,000
	3,000	6	132,000	44		15,000
	0·75	6	3,650	4,880		11,300
	188	6	33,000	176		11,300
String or Einthoven	12	0·002	0·1			1,000,000
	12	0·01	2·5			1,000,000
	125	0·0035	0·1			200,000
	125	0·02	3·3			200,000
	6,000	0·0036	2			1,500,000
	6,000	0·008	50			1,500,000

galvanometers which have been described, in order to give some idea of how their sensitivities compare for current, electromotive force, and quantity of charge. The comparisons are made only for galvanometers having optical projection systems to magnify their deflections.

Moving coil pointer galvanometers have scale deflections of the order of 15 mm. per microamp., and unipivots have much smaller sensitivities.

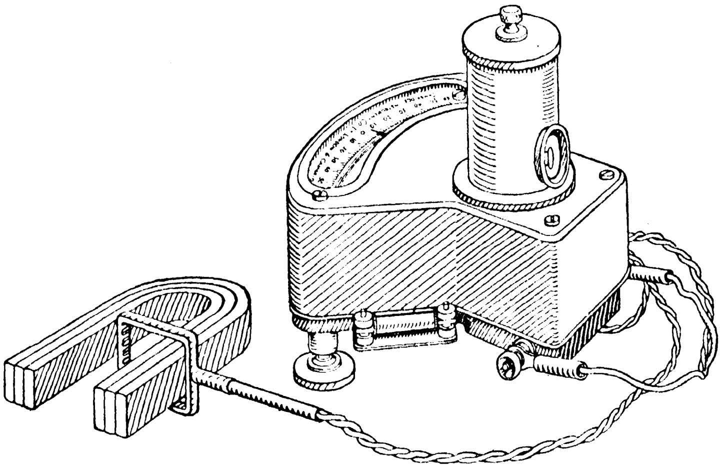


FIG. 204.

The Grassot Fluxmeter

This interesting and most useful instrument may be regarded as a special kind of ballistic galvanometer. It measures the total quantity of charge passing, but has the great advantage of not requiring the charge to pass almost instantaneously. It is, in fact, within reasonable limits, practically independent of the rate at which the charge passes.

In any moving coil instrument used ballistically, the forces tending to bring it to rest are :—

(*a*) The restoring couple of the suspension. This is proportional to the angle turned through by the coil.

(*b*) The electromagnetic damping couple due to the movement of the coil in the field between the magnets. This is proportional to the angular velocity of the coil.

(*c*) The air and frictional damping. This is roughly proportional to the angular velocity.

In the moving coil galvanometer used ballistically, (*a*) is large, (*b*) is small, and (*c*) is negligible. In the fluxmeter (*a*) is very small and (*c*) more negligible than before. Practically the whole control is due to (*b*), the electromagnetic couple.

The general appearance of the fluxmeter will be seen in Fig. 204, which shows the instrument connected to an open wound search coil. The moving coil of the instrument carries a pointer, and is suspended by a silk fibre, the upper end of which is attached to a flat spiral spring A (see Fig. 205) to avoid damage by shock. Current is led in and out of the coil B by means of spirals of thin silver strip CC, acting in opposition, so that the torsional controlling force on the coil is kept to a minimum. On the other hand, the electromagnetic damping is made so great, in proportion, that the only appreciable force which controls the deflections of the coil is that due to this cause. As the field in which the coil moves is due to a powerful permanent magnet, the instrument is practically unaffected by external magnetic fields. The moving system is enclosed in a dust-tight robust metal case fitted with levelling screws and two side spirit levels. A clamping device secures the coil during transit.

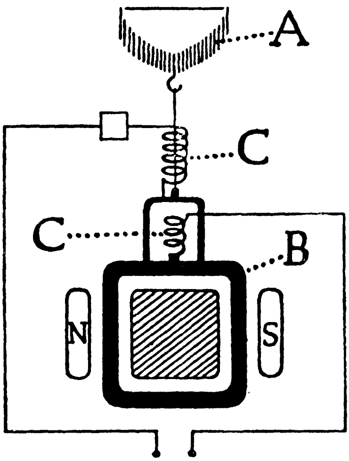


FIG. 205.

The scale of the instrument is approximately 150 millimetres long: the zero is usually at the mid-point of the scale, which is divided into 120 equal divisions, 60 on either side of zero. Each of these divisions represents a definite number of Maxwell-turns, this number (generally about 15,000) being marked on the scale. Thus the total number of lines cut by each turn of the search coil may be obtained by dividing the

scale reading by the number of turns, and if the mean area of the search coil is known, it is easy to calculate the field strength in lines per square centimetre. The search coil is shown looped over one arm of the horseshoe magnet in Fig. 204. It is simply a coil of accurately known area-turns. Full scale deflection corresponds to $1{\cdot}8 \times 10^6$ Maxwell-turns, and the range may be extended, if required, to ten times this value, by shunting the search coil with a non-inductive resistance. It is then possible, with a single-turn search coil, to measure total fluxes up to 18 megalines. On the other hand, very small flux densities may be measured by using a large area multi-turn search coil.

In addition to the pointer, the instrument may be fitted with a plane mirror 5×4 millimetres, and with a lens suitable for a scale distance of one metre, so that for very accurate work a light spot may be used as an index, the beam of light entering through the lens fitted in the back of the case. When used in this way, a given pointer deflection is multiplied about twenty times. Since the coil is not wound on a closed metal former, the instrument may be used undamped on a high resistance circuit as a ballistic galvanometer in the ordinary way, in which case a deflection of one scale division on the instrument is produced by about 0·1 microcoulomb, corresponding to a movement of the light spot of about 25 mm. on a scale at one metre distance.

On open circuit, when the instrument is practically undamped, the free period is approximately 50 seconds. It is aperiodic when connected to a low resistance coil. The resistance of the instrument is approximately 20 ohms, and the search coil may have any value below about 15 ohms; single-turn low resistance search coils may thus be used, in which case the instrument will indicate directly the flux value.

Actual fluxmeters are usually calibrated so that the value of each scale-division in Maxwell-turns is known directly. Thus by winding a search-coil suitable for the particular problem values of flux or intensity varying over a very wide range can be measured.

The obvious uses of the fluxmeter are in measuring the strength of a magnetic field, or the pole-strength of a magnet, and in plotting a B–H curve for iron. It has also many applications in electrical engineering.

*

CHAPTER VII

ELECTROMETERS

Types—Relative Sensitivities to Potential—Relative Sensitivities to Current—The Tilted Gold-leaf Electrometer—Calibration of Vertical Gold-leaf Electroscopes—Calibration of Tilted Gold-leaf Electrometer—Experiments with the Quadrant Electrometer—Natural Leak of Electrometer—Approximate Comparison of High Resistances—Quantitative Measurement of High Resistances—Capacity of Electrometer—Measurement of Dielectric Constants—Measurement of Ionization Currents—The Compton Electrometer—The Lindemann Electrometer—The String Electrometer—Mathematical Theory of the Quadrant Electrometer.

Types

THERE are five kinds of electrometer that are useful for accurate quantitative work. These are—

1. The Tilted Gold-leaf Electrometer (sometimes called the Tilted Electroscope).
2. The Dolezalek Quadrant Electrometer.
3. The Compton Quadrant Electrometer.
4. The Lindemann Electrometer.
5. The String Electrometer.

The Tilted Electrometer is used chiefly for measurements of ionization currents in Radioactivity. The Dolezalek Electrometer is used for instructional work, and for alternating-current measurements. The Compton Electrometer is an improved form of the Dolezalek, in that it is smaller, more sensitive, and able to be controlled in sensitivity. The Lindemann Electrometer is distinguished by the extraordinary rapidity with which it takes up its final reading. It takes considerably less than 1 second to do this. It is used with a photo-electric cell to measure very faint light-intensities, such as those from stars, nebulæ, or comets. The String Electrometer is used to record the intermittent arrival of small charges; particularly for counting α-particles from radioactive substances.

Relative Sensitivities to Potential

Since these electrometers are in general used for different purposes, it is difficult to get any useful comparison of their sensitivities. Some idea of their sensitivities may, however, be gathered from the following table, which, though obviously not exact in detail, gives figures of the right order of magnitude.

Instrument.	Volts on Plate or Needle.	Eyepiece Divisions per volt on Leaf or Quadrants.	Millimetres on scale at 100 cm., per volt on Quadrant.	Time to get Steady Reading.
Tilted Gold-leaf Electrometer.	200	30–200		Less than 1″
Dolezalek Quadrant Electrometer	100		300–8,000	60″–200″
Compton Quadrant Electrometer	50		3,000–30,000	9″–90″
Lindemann Electrometer	40	40–140		Less than 1″
String Electrometer	25	30–60		Less than 1″

Relative Sensitivities to Current

In this table only the sensitivity to potential has been considered; but when ionization currents are being measured it does not follow that if one instrument is more sensitive than another to potential it will also be more sensitive to current. The capacity of the instrument must also be considered. In our table, for example, if the Gold-leaf Electrometer has a sensitivity of 30 scale divisions per volt, and the Dolezalek Electrometer 300 mm. per volt at 100 cm., we may say that for practical purposes the Dolezalek is about ten times as sensitive as the Gold-leaf. [This assumes that a mm. at 100 cm. on a scale is equivalent to 1 scale division, which probably flatters the Dolezalek.]

The capacity of the Dolezalek, however, is about 50 e.s.u. (50 cm.), while the capacity of the Gold-leaf is about 2 e.s.u. Let us suppose that a rate of change of deflection of 1 division (or 1 mm.) per 10 seconds can be measured in each case. The smallest value of $\frac{dV}{dt}$ which can be observed is then $\frac{1}{3000}$ volt for the Dolezalek, and $\frac{1}{300}$ volt for the Gold-leaf. [For the Dolezalek, if 1 mm. per 10 sec. can be measured,

and deflection is 300 mm. per volt, clearly smallest value of $\frac{dV}{dt} = \frac{1}{300} \times \frac{1}{10} = \frac{1}{3000}$ volt/sec.]

Now 50 e.s.u. $= \frac{50}{9 \times 10^{11}}$ farads, and 2 e.s.u. $= \frac{2}{9 \times 10^{11}}$ farads. Thus the smallest ionization current, $\frac{cdV}{dt}$, which can be measured by the Dolezalek is $\frac{50}{9 \times 10^{11}} \times \frac{1}{3000} = 2 \times 10^{-14}$ amperes. Whereas the smallest ionization current which can be measured by the Gold-leaf is $\frac{2}{9 \times 10^{11}} \times \frac{1}{300} = 7 \times 10^{-15}$ amperes, about three times as small.

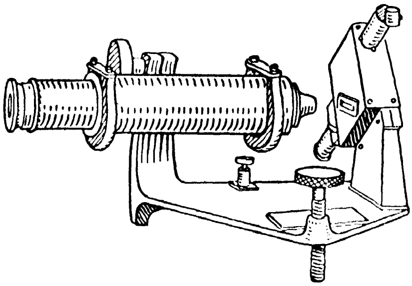


FIG. 206.

The Tilted Gold-leaf Electrometer

The Cambridge Instrument Company's instrument is shown in Fig. 206, and the detailed design of its interior in Fig. 207. The Instrument Company's account of their instrument follows.

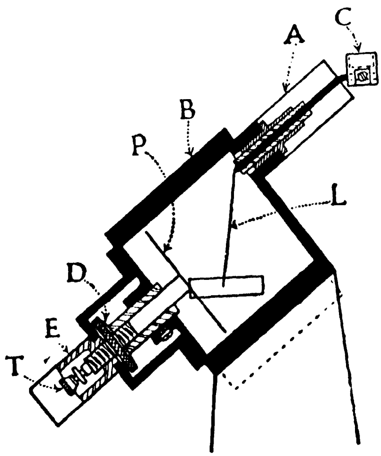


FIG. 207.

The Electrometer illustrated in Fig. 206 was originally designed by Professor C. T. R. Wilson, and subsequently modified by Dr. G. W. C. Kaye. It has a small electrostatic capacity (a few centimetres only), while its sensitivity can be varied within wide limits up to a high

maximum value. The instrument consists of an insulated gold leaf suspended in proximity to an inclined charged plate which attracts it out of the vertical.

It can be shown that, depending on (1) the angle of inclination of the plate, (2) the distance between the plate and the leaf, and (3) the relative potentials of the plate and the leaf, the equilibrium of the leaf may either be stable over the whole of its movement or unstable over part of it. Fig. 208 shows three curves illustrating the possible cases. Of these, curve I

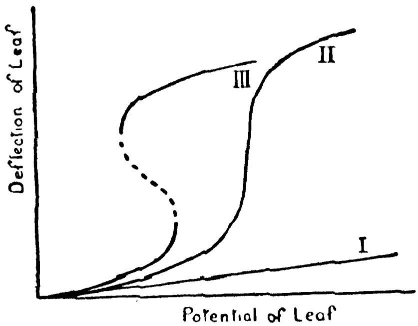


FIG. 208.

shows complete stability and small sensitivity, while curve III combines high sensitivity locally with instability over the dotted region.

The Tilted Electrometer is designed to facilitate the adjustment of the conditions so that an intermediate curve may be obtained, such as curve II, which combines high sensitivity over part of its path with stability throughout. There is, of course, a critical curve which shows infinite sensitivity over a part of its path.

In the diagram, Fig. 207, L is the gold leaf, which is supported by means of a wire passing through a quartz insulating tube to the make and break cup C, and P is the plate which can be adjusted by rotating the ebonite nut D. A is an earthed tube for shielding purposes. The gold leaf is about 1 mm. wide and between 2·9 and 3·4 cm. long. The plate P is connected through the terminal T to a source of steady potential (about 200 volts), an ebonite sleeve E over the terminal serving as a

protection for the observer when adjusting D with the potential on the plate. The rectangular case B of the instrument is earthed, and is provided with windows through which the gold leaf can be observed by means of a microscope fitted with a micrometer eyepiece scale. The complete instrument is mounted together with the microscope upon a rigid base which can be tilted by a supporting screw.

For any particular tilt of the Electrometer there is a plate potential which will yield a maximum sensitivity. The less the tilt, the greater is the plate potential required and the greater the maximum sensitivity, until the critical tilt is reached, beyond which comes instability over some region (*i.e.*, there are two zeros, as in curve III, Fig. 208). The extent of this unstable region increases as the tilt diminishes. By suitable adjustment, a sensitivity of 200 eyepiece divisions per volt over a limited range can be obtained. Except at moderate sensitivities, the deflection is not in general proportional to the potential of the leaf, so that unless the measurements are only relative, and the same part of the scale of the microscope is always used, it is necessary to calibrate the scale either at the beginning or at the end of a set of observations.

Calibration of Vertical Gold-leaf Electroscopes

1. *Natural leak.* The electroscope may be charged by an electrified rod of ebonite or sealing-wax, but this method has the disadvantage of giving a different charge each time. It is better to use a battery of about 300 volts, which may be arranged in the form of a " charging-stick " since no current is taken from it. The original deflection is noted, and the rate of fall of the leaf observed for about a quarter of an hour. If the leak is not greater than about $1\frac{1}{2}$ or 2 scale-divisions in this time the insulation may be regarded as satisfactory. Otherwise the insulation should be scraped or otherwise cleaned.[1]

2. *Calibration of the scale.* Some rather weak but absolutely constant source of ionization, such as uranium oxide, should be brought near the electroscope and placed at such a distance that a rate of fall of the leaf of about 10 divisions per minute is obtained. The leaf should then be timed past each division with a stop-watch, and a graph of deflection against time

[1] Caution here. Cleaning is an expert job. The unskilled should not go further than removing moisture by dry warmth.

should be plotted. This graph will not be quite a straight line, so that in general only a particular range of the scale, specially calibrated, should be used.

When the instrument is calibrated the ionizing power of other agents can be directly compared with that of the source. The sensitivity can be varied by having a small variable condenser (with air between its plates) in parallel with the electroscope.

Calibration of Tilted Gold-leaf Electrometer

The case is earthed and tilted through about 30° at first. The plate is given a potential of about 200 volts, and a small variable potential is applied to the leaf by means of a potentiometer. By varying this potential a curve of potential against deflection may be plotted. This only holds for the particular values used of tilt and plate potential. Either may be varied. For any given value of tilt a particular value of plate potential giving maximum sensitivity may be found. This "critical" value is greater for smaller angles of tilt. If the tilt is gradually diminished the leaf will be found to become unstable for a particular tilt, called the "critical inclination." The greatest sensitiveness is obtained when the tilt is slightly bigger than this critical value.

As the chief use of electroscopes is for radioactive measurements, detailed accounts of the best methods of using them are to be found in standard text-books of radioactivity and radioactive measurements.

Experiments with the Quadrant Electrometer

The Dolezalek Electrometer has been described in detail in Chap. IV, Part I, and an account appears there of its adjustment and use for measurement of direct and alternating potentials, and for comparing capacities. Accounts of its use for a number of other measurements follow.

Natural Leak of Electrometer

If, when the electrometer is used heterostatically, one pair of quadrants is insulated after being charged, it will be found that the charge gradually leaks away. This leak is called the "Natural Leak" of the electrometer.

It is due mainly to three causes. There are always a few ions in the air, and so the air is always very slightly conducting. [But dry air is the best insulator we have.] The insulators on which the quadrants and their connections are supported are not perfect. Their surfaces may be dusty or otherwise dirty.

It is, of course, assumed that the quadrants are charged by having their terminal touched by a wire at the required potential, *not* by having a plug pulled out, which breaks (or appears to break) their connection with the charging battery. The insulation of such a plug is rarely good enough.

The best we can do to reduce the leak is to see that the air is warm and dry near the electrometer, and that the insulators holding the quadrants and their terminals are as clean as possible.

The best way to test the leak is to arrange V so that maximum deflection in either direction is obtained for a potential of 1 volt. Then if the deflection, when the quadrants are insulated, does not diminish faster than 1% per minute the leak is not too large. If it is much larger something must be done about it.

A very odd trouble may arise.

If the quadrants are left insulated but uncharged, a deflection may gradually appear and mount up. This is probably due to the leakage of charge which has got somehow on to the insulators. If a little radium cannot be obtained to ionize the air and discharge the insulators (and it is unlikely that it can), the charge may be perhaps removed by cleaning.[1]

Approximate Comparison of High Resistances

If the quadrants are given a charge which causes maximum deflection, they can be made to discharge to earth through various resistances. The values of these resistances will be roughly proportional to the times taken by the deflection to decrease from one definite value to another, provided that these values are not too far apart. If the maximum deflection is 25 cm., the time taken for the deflection to go from 20 cm. to 15 cm. would be fairly suitable. It is no use using this method if the rate of fall is more than about 1 cm./sec.

This method is actually very rough, and theoretically wrong (though not very far wrong if the change of deflection is small

[1] All switching should be made in clean paraffin wax of good quality. Wiring should be screened when possible.

and the time great), but it is extremely easy and convenient for a quick test.

Values of the specific resistance of various good insulators (according to Kaye and Laby's Physical Constants) are as follows :—

Paraffin wax	3×10^{18} ohms/cm. cube
Mica	9×10^{15} ,, ,,
Sulphur	4×10^{15} ,, ,,
Porcelain, Ebonite	2×10^{15} ,, ,,
Quartz	$1{\cdot}2 \times 10^{14}$,, ,,
Jena glass	2×10^{14} ,, ,,
Soda-lime glass	5×10^{11} ,, ,,

In practice paraffin wax is apt to be impure and have a relatively low resistance. Mica will not accommodate itself to the shape one wants, and neither will porcelain. Ebonite is very apt to be impure, and so is glass; so that sulphur, quartz, and air are the best practical insulators. Sulphur may be fused and run into a mould, but if its temperature is allowed to rise enough to change its colour from light yellow to orange it loses its insulating properties. Really good ebonite is satisfactory, but its surface may be made conducting by exposure to light, and the only cure is to remove that surface.

It is sometimes instructive to test the ebonite sold for panels for wireless sets, and to find the " natural leak " across many small condensers. In fact many of the smaller and cheaper wireless parts are worth investigating.

Quantitative Measurement of High Resistances

The circuit is connected up as in Fig. 209. B is a cell or accumulator, C a condenser whose capacity depends on the size of the resistance R being measured. W is the water-resistance, and H the high-tension battery, which should have such a value that B gives full-scale deflection. K_1 and K_2 are keys made by putting mercury in small holes cut in *good* paraffin wax. Contact between two holes is made with a piece of copper wire fixed into an insulating handle of sealing-wax.

First charge the condenser C to potential B by closing K_1 with K_2 open. Observe the deflection d_1.

Close K_2 by joining 1 to 3. Open K_1, starting a stop-watch as you do it.

The deflection will decrease as the charge leaks to earth through R. Take the time t required for the deflection to fall to about half its original value. Let the new deflection be d_2.

Then it can be shown that

$$R = \frac{t}{C \log_e\left(\frac{d_1}{d_2}\right)}$$
$$= \frac{0{\cdot}4343t}{C \log_{10}\left(\frac{d_1}{d_2}\right)}$$

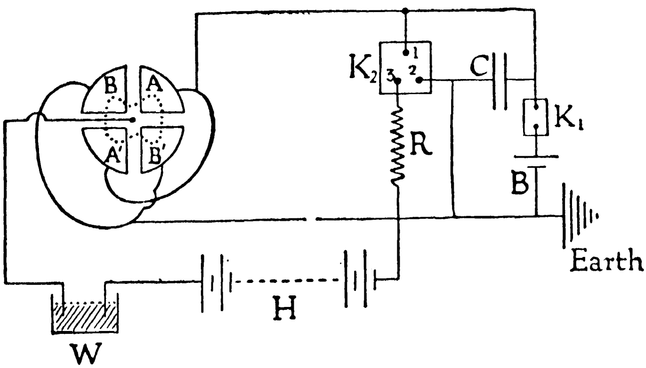


FIG. 209.

if t is measured in seconds, C in farads, and R in ohms. Several times should be observed, and there should be no systematic variation of

$$\frac{t}{\log\left(\frac{d_1}{d_2}\right)}.$$

This method is useful for finding the specific resistance of a liquid. The liquid can be put in a long cylindrical tube with metal ends.

Capacity of Electrometer

Connect up the apparatus as in Fig. 210. The circuit is the same as for Fig. 209, except that R is removed, the capacity

C is much smaller, and the connections of K_2 have been changed. C should have a capacity of about 80 cm., which should be accurately known.

A guard-ring condenser is the easiest to get both small and

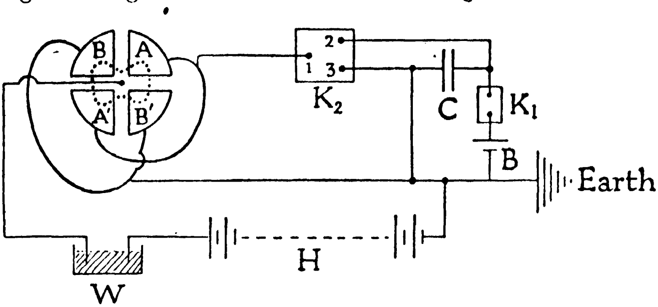


FIG. 210.

accurate. Such a condenser is roughly sketched in Fig. 211. The upper plate, which is earthed, is movable, and a scale gives an accurate reading of the distance between the plates. The lower (insulated) plate is supported on good ebonite and surrounded by a guard-ring.

The guard-ring is connected to the plate when the condenser is given its charge by having a potential put on the insulated plate, but in all later operations the guard-ring is left insulated.

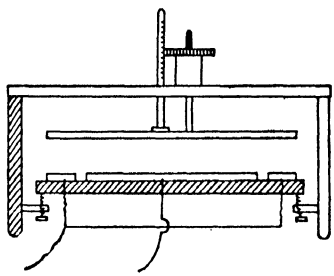


FIG. 211.

Close K_1, and close K_2 from 1 to 2. This charges up C and the electrometer together, and the electrometer is deflected an amount d_1. Disconnect K_2, and reconnect it from 1 to 3. This discharges the electrometer to earth, and it should return to zero.

Open K_1, and the charged plate of the condenser is insulated and separated from its guard-ring.

Switch K_2 back so that it connects 1 and 2. The condenser C shares its charge with the electrometer, and a smaller deflection d_2 should result.

Then if C_e is the electrometer capacity, it is easy to show that

$$\frac{d_1}{d_2} = \frac{\text{potential of charge on C}}{\text{potential of same charge on } (C + C_e)}$$
$$= \frac{C + C_e}{C}.$$

So $$C_e = C \cdot \frac{d_1 - d_2}{d_2}.$$

If C is large compared with C_e we may share the charge a number of times by changing the key K_2 backwards and forwards between the two positions given. Each time the potential is reduced in the ratio $\frac{C}{C + C_e}$. It can be shown that if this is done n times, so that the final deflection is d_{n+1}, then

$$C_e = C\left\{\sqrt[n]{\frac{d_1}{d_{n+1}}} - 1\right\}.$$

The capacity of the guard-ring condenser is, of course, $\frac{A}{4\pi d}$ e.s. units, if A is the area of the insulated plate. This is equal to $\frac{A}{9 \times 10^{20} \times 4\pi d}$ e.m. units, or $\frac{A}{9 \times 10^{11} \times 4\pi d}$ farads, or $\frac{A}{9 \times 10^{5} \times 4\pi d}$ microfarads.

The distance d between the plates is measured in cm. on the scale, but unfortunately it is rather difficult to find the zero of the scale. The best way is to plot the observed value of d on the scale against the reciprocal of the capacity of the electrometer found by comparison with a standard by the method of Experiment 4. The resulting graph will probably be a straight line, and the value of d where the graph crosses the axis is the effective zero.

When the capacity of the electrometer is known it can be used to find the capacities of other small condensers by sharing the charge. It is very interesting to have in parallel with the electrometer a condenser consisting merely of two large metal plates, one insulated by being stood on legs of sulphur or paraffin wax, and the other earthed. If one moves the earthed plate about, varying the distance between the plates, and

the area of one plate opposite the other, one gets a very good idea of the effect of the relative position of the plates on capacity, and (what is now clearly seen to be a closely related phenomenon) the effect of the proximity of an earthed conductor on the potential of an insulated charged conductor.

This most instructive qualitative experiment should on no account be omitted.

Measurement of Dielectric Constants

If a slab of dielectric of constant k and thickness t cm. is put between plates of a parallel-plate condenser d cm. apart, having a surface density of charge σ, a unit charge being brought from the earthed plate to the insulated one is only brought against an intensity of $\frac{4\pi\sigma}{k}$ in the dielectric.

So the total work done to bring the charge from one plate to the other is $\frac{4\pi\sigma}{k} \cdot t + 4\pi\sigma(d - t)$, instead of being $4\pi\sigma d$ as it was before the slab of dielectric was introduced.

Thus the potential of the insulated plate is reduced in the ratio

$$\frac{\frac{4\pi\sigma t}{k} + 4\pi\sigma(d - t)}{4\pi\sigma d}$$

or

$$\frac{d - t + \frac{t}{k}}{d}.$$

So the capacity of the condenser is increased in the ratio

$$\frac{d}{d - t + \frac{t}{k}}.$$

Thus by finding with our electrometer the proportional change in the capacity of a condenser due to the introduction of a slab of dielectric we can calculate the dielectric constant.

Suppose, for example, t was chosen equal to $\frac{d}{2}$, and the introduction of the dielectric made the capacity increase in the ratio $\frac{3}{2}$.

Then
$$\frac{d}{d - \frac{d}{2} + \frac{d}{2k}} = \frac{3}{2}.$$
$$\frac{2}{1 + \frac{1}{k}} = \frac{3}{2}.$$
$$4 = 3 + \frac{3}{k}.$$
$$\therefore \quad k = 3.$$

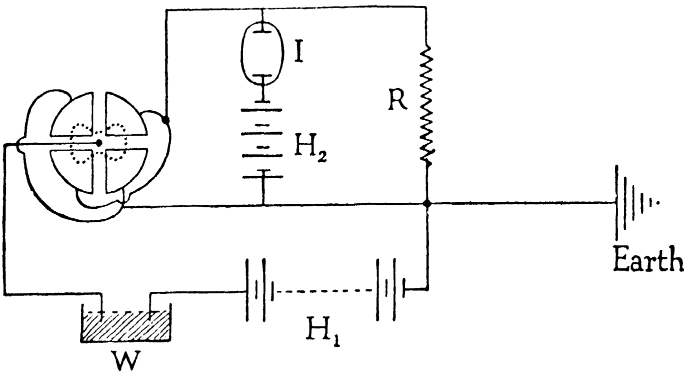


FIG. 212.

Measurement of Ionization Currents

One of the principal uses of the quadrant electrometer has been to measure the very small currents occurring in ionization experiments.

The most direct method is Bronson's (see Fig. 212). The battery H_2 drives a current through I, the ionization chamber, and through R, a known high resistance. The electrometer has first been calibrated so that the relation between its deflection and the potential difference of the quadrants is known.

When the battery H_2 is put in the circuit the current may take some time to settle down to its steady state. When it has settled down the potential difference V across the resistance R will be known. So the ionization current flowing will be $\frac{V}{R}$. The resistance R may have to be very high—perhaps as much as 10,000 megohms.

A very accurate method is Townsend's (Fig. 213). The key K_2 is opened, and K_1 is shut, and a stop-watch is started. Current then begins to flow through the ionization chamber I and charge up the condenser C. The electrometer begins to show a deflection. It is brought back to zero and kept there approximately by adjusting the variable resistance of the potentiometer R. After a reasonably measurable time t has passed K_1 is opened, and the deflection is brought exactly to zero by adjusting R. The fall of potential across the condenser C is then equal to the rise along the resistance R from earth. We can find this potential V by knowing the values of the subdivisions of R and by knowing V.

The charge on C is now VC.

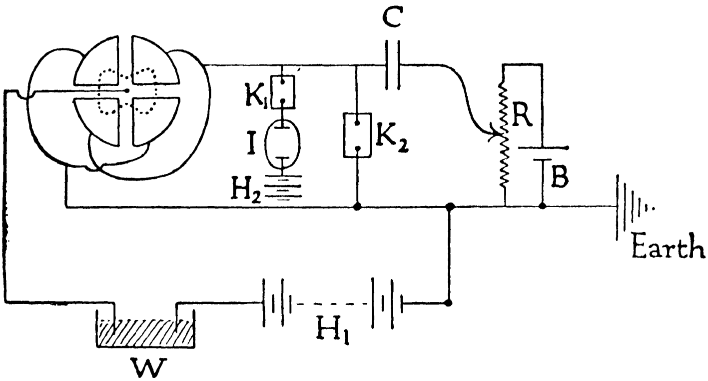


FIG. 213.

This charge flowed into the condenser through the ionization chamber in time t. C, V, and t are known. Then the ionization current i is clearly given by

$$i = \frac{VC}{t}.$$

The Compton Electrometer

The Compton Quadrant Electrometer (Fig. 214) differs from the Dolezalek in two ways. The movement is smaller, and it contains an arrangement by which the needle may be tilted at will, and one of the quadrants may be moved vertically with respect to the other three. It can be shown that the sensitivity

can be altered either by tilting the needle or by raising or lowering one quadrant. The sensitivity can thus be controlled from outside without disturbing the instrument. Its capacity is between $\frac{1}{5}$ and $\frac{1}{10}$ of that of the Dolezalek, and its sensitivity is considerably greater.

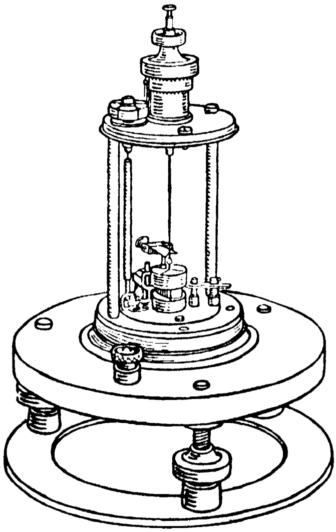


FIG. 214.

The Lindemann Electrometer

The Lindemann Electrometer is a most remarkable instrument. Its general principle is that of the Quadrant Electrometer, but its dimensions are 9 × 4 × 2·5 cm., and its weight less than 3 ounces. It requires no levelling, is very robust, takes up its final reading in less than 1 second, and has a capacity as small as, and possibly smaller than, that of the Gold-leaf Electrometer.

In connection with photo-electric measurements of the intensity of light received from stars, nebulæ, comets, etc., it is desirable to mount an electrometer and a photo-electric

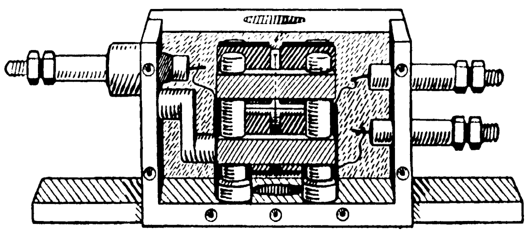


FIG. 215.

cell just behind the focus of an equatorial telescope. It is sometimes an advantage to mount the electrometer on the telescope, and it is then necessary to use a compact electrometer the sensitivity and zero of which are unaffected by the movement of the telescope. The Electrometer illustrated in Fig. 211

was designed by Prof. F. A. Lindemann to meet these requirements. It is compact and robust, has a high sensitivity, a stable zero, and does not require levelling. Since the capacity is less than two centimetres, and the period less than one second, the instrument may be conveniently used as a laboratory electrometer. With a suitable leak the Electrometer forms a convenient quick-period galvanometer that can be usefully applied to laboratory measurements. The instrument is similar in principle to the quadrant electrometer. It consists of a needle, which in this instance really has the shape of a needle, suspended at its centre so that it can rotate between four cross-connected plates which replace the ordinary quadrants. An important feature of the instrument is the mounting of the torsion fibre, which is fixed at both ends under tension so that the centre of rotation of the needle is fixed; the rotation can therefore be determined by observing the movement of one end of the needle through a microscope. This renders a mirror unnecessary, and the consequent small inertia of the moving parts enables a small torsional restoring force to be used without unduly increasing the period. The torsion fibre is of gilded quartz about 0·004 to 0·006 millimetre in diameter and 1·4 centimetres long. It is stretched under a tension of about half its breaking strain on a U-shaped frame of quartz rod, 3 millimetres in diameter, the ends of the fibre being fixed in saw-slits about 1 millimetre deep in the extremities of the frame. As the frame and the fibre are both of quartz, no change of tension occurs when the temperature of the instrument changes. In order to facilitate balancing, the needle is made of two parallel fibres of gilded glass (0·02 millimetre diameter, 2 centimetres long), mounted one on either side of the torsion fibre, with their ends and centres sealed together. The quadrants comprise four plates, about 1·5 centimetres broad and 1 centimetre high, through which slots 2 millimetres wide are cut for the passage of the needle. They are mounted about 5 millimetres apart on quartz rods, the moving system being mounted between them so that the centre of rotation of the needle coincides with the centre of symmetry of the plates.

The instrument is enclosed in a light metal box having apertures in the centre at the top and bottom, closed with

microscope cover slides. The leads from the two sets of plates pass through quartz tubes to terminals at one end, while the lead from the needle is carried out in a similar way to a terminal at the other end. An earthing terminal enables the effect of stray electrostatic fields to be eliminated. A moisture absorption chamber is fitted, into which a small quantity of phosphorus pentoxide is introduced as a drying agent, enabling the Electrometer to be used under humid conditions.

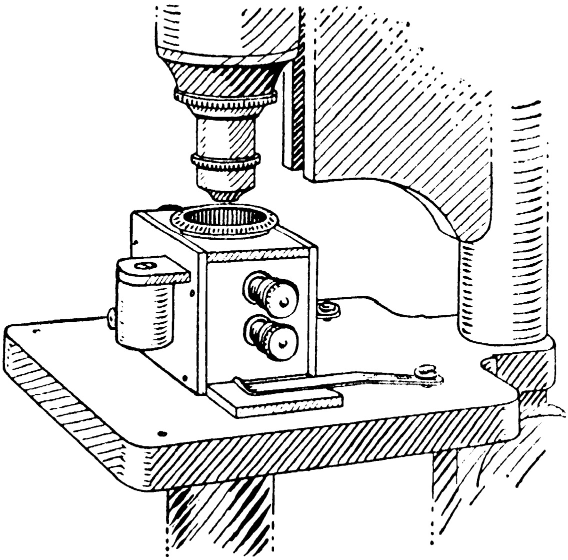


FIG. 216.

For general use, the box can be placed upon an ordinary microscope stand, and the upper end of the needle observed, illumination being obtained in the usual way through the window underneath the box (see Fig. 216). The sensitivity can be adjusted by altering the potentials on the plates. The ultimate sensitivity depends also upon the power of the microscope used, and since a considerable magnification can be employed without serious loss of definition, a sensitivity of considerably more than 400 eyepiece divisions per volt may generally be obtained. The Electrometer can also be used in conjunction with a photographic recording camera.

The String Electrometer

This electrometer is designed on the same principle as the Einthoven Galvanometer, but an electric instead of a magnetic field causes the deflection. The Cambridge Instrument Company's account follows.

The String Electrometer shown in Fig. 217 is of a form suggested by Prof. Einthoven. It has small capacity (approximately two centimetres), short period, is deadbeat, and is therefore suitable for following and, when used with a suitable camera, for recording, rapid changes in potential, particularly

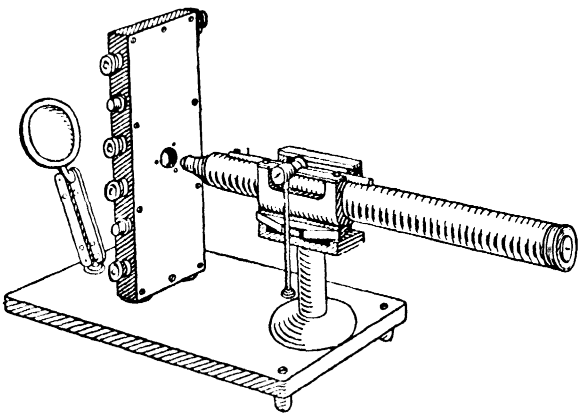


FIG. 217.

in connection with the study of discontinuous phenomena. The sensitivity of the instrument compares well with Gold-leaf Electrometers, a force of the order of 10^{-8} dyne being sufficient to deflect the fibre. The moving system is a conducting fibre held under tension parallel to two metal plates, between which a constant potential difference is maintained. The potential to be measured is connected to the fibre, and the consequent deflections are either observed by means of a microscope or recorded by means of a camera. The fibre is easily replaced. For low voltage work the fibre is of plated quartz (0·002 to 0·003 millimetre in diameter) and is mounted on an invar rod to compensate for temperature variations, while for high voltage work it is of copper (approximately 0·011 millimetre

in diameter) mounted on a copper rod. The metal plates can be adjusted laterally to set the electrical centre of the system coincident with the fibre, the deflections of which in either direction are then equal for equal positive or negative potentials, and vary with the potential being measured. By adjusting the tension in the fibre, or by altering the potential applied to the plates, the sensitivity can be varied over a wide range. At the suggestion of Prof. F. A. Lindemann, the instrument is designed as a separate unit that can be conveniently assembled with other apparatus or used on a microscope stage. The moving system is entirely enclosed within a narrow rectangular case, which is earthed, thus effectively shielding the fibre from draughts and from the effect of external electrostatic fields.

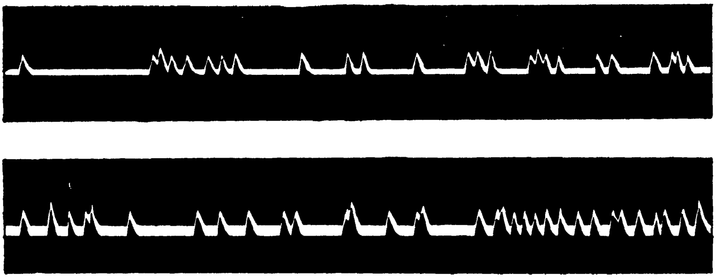


FIG. 218.

The case is provided with a small window so arranged that the microscope objective can be brought into close proximity to the fibre. A microscope, with micrometer eyepiece scale graduated in tenths of a millimetre, is supplied for reading the fibre deflections. The law governing the deflections of the string is not linear, the sensitivity being much higher for small potentials of the order of 0·01 volt than for potential differences of about one volt.

The complete instrument, with microscope and illuminating mirror, is mounted on a rigid stand. An optical condenser can be fitted, by means of which the fibre can be sufficiently illuminated to enable its image to be projected on a screen.

Fig. 218 shows two interesting records which were obtained by Rutherford and Geiger with an earlier form of the String

Electrometer, in connection with their electrical method of determining the number of α particles in a given space. They show the rate at which α particles from radium enter a partially exhausted vessel exposed to an electrostatic field (1,500 volts). The rays ionize the air in the vessel, the ions produced being multiplied several thousand times by collision, and the effect of a single α particle is thus sufficiently magnified to give a detectable deflection on the String Electrometer. The records were taken photographically on a film moved at a rate of 180 centimetres per minute, the upper record showing the entry of particles at the rate of about 600 per minute and the lower record showing entries at about 900 per minute.

Mathematical Theory of the Quadrant Electrometer

Let AA′ be one pair of quadrants, BB′ the other pair (Fig. 219).

Let N be the needle and θ the angle through which it is deflected from the zero position.

Let the potential of AA′ be v, of BB′ zero.

Let V be the potential of the needle.

Let the radius of the needle be r, and let it be hung midway between the quadrants, at a distance t from each. When the needle turns through an angle θ in the direction of BB′, an area $2r^2\theta$ of needle is transferred from the AA′ to BB′. [θ is of course measured in radians.]

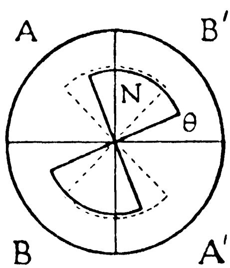


FIG. 219.

The AN condenser is thus decreased in capacity by an amount $\frac{2r^2\theta}{4\pi t}$, or $\frac{r^2\theta}{2\pi t}$, and the BN condenser is increased by an equal amount.

Thus a charge $\frac{r^2\theta}{2\pi t}(V - v)$ has been lost by the AN condenser, and a charge $\frac{r^2\theta V}{2\pi t}$ gained by the BN condenser.

$$\therefore \text{ Electrical energy lost by AN} = \tfrac{1}{2}\left[\frac{r^2\theta}{2\pi t}(V - v)\right] \times [V - v]$$
$$= \frac{r^2\theta(V - v)^2}{4\pi t}.$$

Similarly energy gained by BN $= \dfrac{r^2\theta V^2}{4\pi t}$.

$\therefore$ Energy supplied to electrometer to produce deflection θ

$$= \frac{r^2\theta}{4\pi t}[V^2 - (V - v)^2]$$
$$= \frac{r^2\theta}{4\pi t}(2Vv - v^2)$$
$$= \frac{r^2\theta v}{2\pi t}\left(V - \frac{v}{2}\right).$$

This energy is proportional to θ. As the work done by a couple is the product of the couple and the angle through which it turns,[1] it follows that

$$\text{Electrostatic couple} = \frac{r^2 v}{2\pi t}\left(V - \frac{v}{2}\right).$$

But in the equilibrium position this is equal to the restoring couple of the suspension. Let this be C per radian twist.

Then if θ is now taken as the final steady value of the deflection in equilibrium,

$$c\theta = \text{electrostatic couple}$$
$$= \frac{r^2 v}{2\pi t}\left(V - \frac{v}{2}\right).$$
$$\therefore \quad \theta = \frac{r^2 v}{2\pi c t}\left(V - \frac{v}{2}\right).$$

Since v is in general small compared with V, we may write with approximate accuracy—

$$\theta = KVv,$$

where K is a constant of the electrometer.

This is the rough equation when electrometer is used heterostatically. When it is used idiostatically (see p. 58), $V = v$, and

$$\theta = KV^2.$$

If, however, V is made large when the electrometer is used heterostatically it is found that the sensitivity does not continue to increase in proportion to V. It reaches a maximum when V has some value of the order of 500 volts for a Dolezalek electrometer.

This is explained by the occurrence of another couple which has not yet been mentioned.

[1] Where, as here, the couple is constant. This would not apply to a twisted wire.

When the needle is deflected, and the restoring couple of the suspension acts, there is another restoring couple due to the distortion of the lines of force between the needle and the quadrants. This acts even if $v = 0$, and can be shown to be proportional to V^2 and θ.

Thus instead of $$c\theta = \frac{r^2 v}{2\pi t}\left(V - \frac{v}{2}\right)$$

we should have had

$$c\theta + kV^2\theta = \frac{r^2 v}{2\pi t}\left(V - \frac{v}{2}\right).$$

Thus $$\theta = \frac{r^2 v\left(V - \frac{v}{2}\right)}{2\pi t(c + kV^2)}.$$

If now we again neglect $\frac{v}{2}$ in comparison with V, and choose constants

$$k_1 = \frac{2\pi tc}{r^2}$$
$$k_2 = \frac{2\pi tk}{r^2}$$

our equation becomes

$$\theta = \frac{Vv}{k_1 + k_2 V^2}.$$

By differentiating we can show in the ordinary way that θ is a maximum when

$$V = \sqrt{\frac{k_1}{k_2}}.$$

The equation $\theta = \dfrac{Vv}{k_1 + k_2 V^2}$ is in practice the most useful.

It is obvious that all these methods of attack are full of approximations. A more accurate equation can be shown to be

$$\theta = \frac{Av\left(V - \frac{v}{2}\right) - BV^2}{c + CV^2}$$

where A, B, and C are constants of the electrometer.

Both the last equations have an interesting application in the Compton electrometer.

It can be shown that by tilting the needle we can make the distortional electrostatic couple oppose, instead of supporting, the torsional couple. The equation would then take the approximate form

$$\theta = \frac{\mathrm{V}v}{k_1 - k_2\mathrm{V}^2},$$

and if $\mathrm{V} = \sqrt{\frac{k_1}{k_2}}$ the sensitivity is infinite!

It is for this reason that the Compton electrometer has its sensitivity controlled by tilting the needle; in this way the sensitivity can be made very great. However, the time of swing of the needle increases with the sensitivity; and the sensitivity is thus limited by the necessity of taking one's reading before the end of Time!

*

CHAPTER VIII

ALTERNATING CURRENTS

Alternating Current—Simple Harmonic Motion and Simple Harmonic Current—Uses of Alternating Current—Measurement of Alternating Current—The Duddell Oscillograph—Calibration of Instruments measuring i^2—Thermo-ammeters—Dynamometers—Electrostatic Voltmeters—Cathode-ray Oscillograph—Inductance in A.C. Circuits—Capacity in A.C. Circuits—Reactance of Condenser—Exact Treatment of Condenser Problem—Inductance with Resistance in A.C. Circuits—Capacity with Resistance in A.C. Circuits—Capacity, Inductance and Resistance combined in an A.C. Circuit—Vector Diagrams—Power in A.C. Circuits—Wattless Currents—General Properties of A.C. Circuits—The Series Resonance Circuit—Ten Numerical Examples.

Alternating Current

IMAGINE that a flat coil of wire is rotating about a diameter BC fixed perpendicular to a magnetic field H, as in Fig. 220.

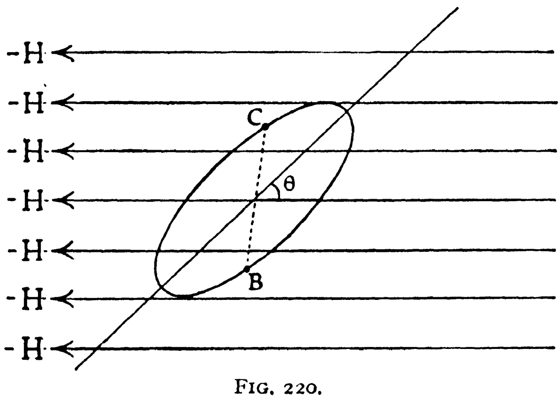


FIG. 220.

If the plane of the coil makes an angle θ with the field at time t, it is clear from the arrangement of the coil that the total number of lines of force, or total flux, threading the coil will

be proportional to $\sin\theta$. No flux will thread it when $\theta = 0$; a maximum when $\theta = \frac{\pi}{2}$.

An E.M.F. will be induced in the coil as it rotates by the change of flux through it. This E.M.F. will, in absolute units, be equal to the rate of change of flux at any instant.

If the coil is rotating at constant speed, the E.M.F. will be proportional to $\frac{d}{dt}(\sin\theta)$, or to $\cos\theta$, since θ increases in proportion to t. We may therefore write

$$E = E_0 \cos\theta$$

where E_0 is the biggest value attained by the E.M.F., and θ is the angle between the plane of the coil and the direction of

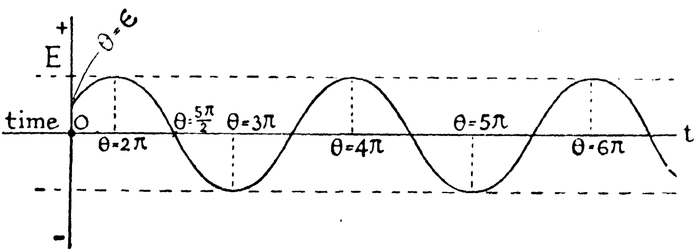


FIG. 221.

the field. E is thus a maximum when $\theta = 0$, and a minimum when $\theta = \frac{\pi}{2}$ or $-\frac{\pi}{2}$, or any odd multiple of these.

Suppose now that our coil was making an angle ε with the plane of the field when we began to reckon t, and suppose it was turning round n times per second. θ would be increasing (in radian measure) at the rate of $2\pi n$ per second, and it would have been ε when t was 0. We can thus write our equation more informatively—

$$E = E_0 \cos(2\pi nt + \varepsilon). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Fig. 221 shows how E behaves as t increases. E is clearly a positive maximum when θ or $\{(2\pi nt + \varepsilon)\}$ is equal to 2π, 4π, 6π, etc.; a negative maximum when θ is π, 3π, 5π, etc., and zero when θ is any odd multiple of $\frac{\pi}{2}$. The most important

feature of the graph is that the value of E alternates in a perfectly regular way, following a cosine-curve as t increases.

If the coil of wire in which this E.M.F. were being induced had no break in it, the E.M.F. would clearly cause an alternating current to flow, and by some such device as that of Fig. 148, p. 263, the system could be used as a very primitive kind of alternating-current dynamo to supply an external circuit.

Simple Harmonic Motion and Simple Harmonic Current

Before we can deal further with this kind of current, we must consider the nature of Simple Harmonic Motion in some detail. In Fig. 222 let a point P move round a circle, centre O and diameter AB, with constant speed. Let the speed be such that OP turns through an angle of p radians per second.

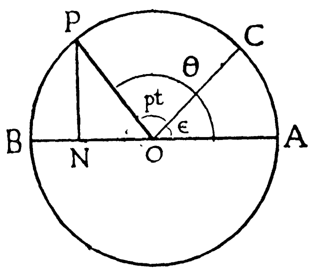


FIG. 222.

Suppose we began to reckon the time t from a particular instant when P was at the point C on the circle. Then, if angle AOC is ε, angle AOP is $(pt + \varepsilon)$ at time t.

The quantity p is known as the *angular frequency* of OP, since it is the number of radians swept out by OP per second. Now let N be the foot of the perpendicular dropped from P upon AB.

Then $\quad \text{ON} = \text{OP} \cos \theta.$

So we may write $\quad x = \alpha \cos (pt + \varepsilon) \quad . \quad . \quad . \quad . \quad . \quad (2)$

if $\quad x = \text{ON}, \text{ and } \alpha = \text{OP}.$

Simple Harmonic Motion is motion of a point which moves so that its acceleration is always directed to a fixed point, and is proportional to its distance from that point.

Now the only acceleration of P is, by the ordinary theory of circular motion, constant, of the value $p^2\alpha$, and directed along the radius PO. The factor p^2 is of course a constant.

Thus the acceleration of N is the resolved part of this acceleration in the direction NO. It is thus $p^2a \cos \theta$. But $a \cos \theta = \text{NO}$, so that the acceleration of N towards O is $p^2 . \text{NO}$, or p^2x. Thus N describes simple harmonic motion. Its

period of complete oscillation is the same as the time P takes to go once round. In going once round OP sweeps out an angle of 2π at p radians per second. So the period is $\frac{2\pi}{p}$ seconds.

Now let us consider alternating current in this connection. Let us suppose for simplicity that the induced current in the coil is directly proportional to the induced E.M.F. E. [In fact it is not, in general; but no harm will be done in the discussion of alternating currents if, for the purpose of this paragraph, we assume that it is.]

Instead of the E.M.F. equation (1), we can write

$$i = i_0 \cos (2\pi nt + \varepsilon) \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where i_0 is now the biggest value the current reaches. The equations (2) and (3) have exactly the same general form, except that p in (2) is replaced by $2\pi n$ in (3).

But p is an *angular* frequency measured in radians per second, whereas n is an *absolute* frequency measured in revolutions per second. Clearly to do 1 revolution per second is the same as to do 2π radians per second; so that the difference is simply one of notation. In fact

$$p = 2\pi n$$

and

$$n = \frac{p}{2\pi}.$$

We may thus write the equation for an alternating current in the form

$$i = i_0 \cos (pt + \varepsilon) \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where this represents a current alternating $\frac{p}{2\pi}$ times per second.

The exact parallelism of equations (2) and (4) gives our reason for calling a current represented by (4) a *simple-harmonic* or *sinusoidal* current.

The state of such a current at any instant, and its general behaviour, can thus be best represented by such a figure as Fig. 223, for in this figure the direction and magnitude of ON at any instant represents the direction and magnitude of the current i. And the length of the radius a represents the maximum value i_0 of the current. This value is called the *amplitude*.

The angle AOP, or θ, which determines the position of N in Fig. 222, and thus the value of the current i at time t, is known as the *phase* or *phase-angle*. Fig. 222 may be adapted as in Fig. 223 to represent a sinusoidal current. ON now represents i, the instantaneous current, OA represents i_0, the maximum current, and θ represents the phase of the current at time t. ϵ represents the phase of the current at time $t = 0$.

This account has dealt with the nature and representation of the sinusoidal alternating current. Any more complicated form of the A.C. current can be represented by adding the effects of two or more sinusoidal currents, since by Fourier's theorem any oscillation whatever can be represented by the addition of a sufficient number of simple-harmonic oscillations. The simple form of dynamo mentioned in the last section is never used in practice, since A.C. dynamos are generally used to generate very high voltages, and a brush system is unsatisfactory for this purpose.

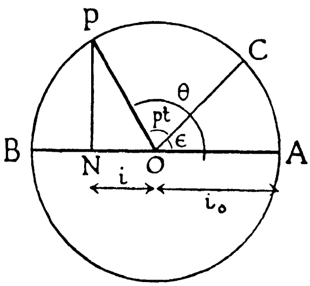


FIG. 223.

A.C. current can, however, be generated by the almost equally simple method of making the field revolve round a stationary armature.

It is impossible here to find room to discuss at all the immense subject of A.C. generators; and indeed an adequate discussion of such problems cannot be found outside text-books of electrical engineering. Here it is only possible to mention the fundamental principle of the production of alternating current, and to continue with other principles relating to it.

Uses of Alternating Current

When electrical power is transmitted over long distances for such purposes as heating, lighting, or driving machinery, the power-loss in transmission is proportional to the product of the resistance and the square of the current. The power transmitted, on the other hand, is proportional to the product of current and voltage.

Thus the higher we raise the transmission voltage for given

power the smaller the current we need carry, and the less the loss of power in the resistance.

It is thus clear that for economy of working long-distance power-transmission should be at the highest practicable voltage.

It is, however, for obvious reasons, impracticable to have voltages high enough to be dangerous for household supply. In practice household voltages are of the order of 200 volts, and long-distance transmission voltages may be over 100,000 volts.

The economical generation voltage of electricity is of the order of 5000 volts; far above the safe value, and far below the transmission value.

Frequent transformation from medium to high voltage, and from high voltage to low voltage, is thus necessary. This cannot be done easily with direct current, but is a comparatively simple matter with alternating current. Alternating current thus provides the only possible scheme for supplying large areas with electrical energy. It is suitable for heating, lighting, and driving machinery but, of course, not suitable for electrolytic work or for charging accumulators. For these purposes, or for any which require direct current, an A.C. supply must be rectified, or made unidirectional, again.

Measurement of Alternating Current

The ordinary moving coil or moving magnet galvanometer has no chance of following the oscillations of an alternating current because of the high moment of inertia of the moving system. It simply shows the mean value of $i_0 \sin (pt + \varepsilon)$, which is, of course, zero by symmetry. Mathematically it is easy to show that the mean current is zero. If Q is the quantity of charge passing during one complete cycle,

$$Q = \int_0^T i\,dt$$

where $T = \frac{2\pi}{p}$, the time of a complete cycle.

Thus
$$Q = \int_0^{\frac{2\pi}{p}} i_0 \cos (pt + \varepsilon)dt$$
$$= i_0 \{\sin (2\pi + \varepsilon) - \sin \varepsilon\}$$
$$= 0.$$

There are, however, a number of ways of measuring alternating current and E.M.F.

Alternating-current measuring instruments belong to two main types:

1. A rapidly-moving instrument which can respond to the alternations of the current. The Vibration Galvanometer, the Duddell Oscillograph, and the Cathode-ray Oscillograph are examples of this type.

2. A slow-moving instrument which responds to the square of the current (and is thus independent of its direction). The hot-wire ammeter which measures the heating effect, the dynamometer which measures the attraction between two coils each carrying the current, and the electrostatic voltmeter (for measuring E.M.F. only) are examples of this type.

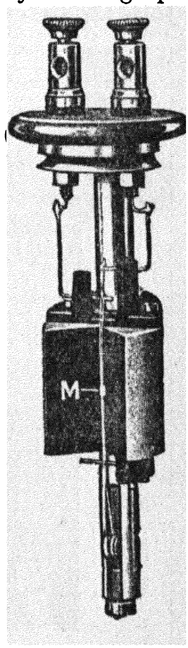


FIG. 224.

The Duddell Oscillograph

The Duddell Oscillograph resembles the Einthoven String Galvanometer in principle, but has two strings[1] instead of one. A mirror is fixed from one string to the other. The current is sent through the strings in series, so that when one moves inwards the other moves outwards. The mirror consequently oscillates with the frequency of the current through the strings, and a reflected beam vibrates over a considerable range on the scale. The Duddell oscillograph is thus in a sense a vibration galvanometer of very high natural frequency. The frequency is made high by stretching the strings tightly. The oscillograph can also be made with an electrostatic controlling field, as in the Einthoven string electrometer. In this form it is itself an electrometer.

The Cambridge Instrument Company's account of their oscillographs follows.

An electromagnetic vibrator is illustrated in Fig. 224, and an electrostatic vibrator for use on circuits above 1000 volts is shown in Fig. 225. The electromagnetic unit is of similar

[1] The "string" is really a flat metal strip.

design to the original Duddell pattern, but is of smaller dimensions. It is mounted in a separate oil bath placed in the gap between the pole pieces of a permanent magnet. A lens is recessed into the oil bath so as to be as near as possible to the mirror M. Provision is made for adjusting the position of the vibrator. The undamped natural period of the vibrator in air is 0·00033 second, in oil it is 0·00053 second, and the D.C. sensitivity (in oil) is 60 milliamperes for a deflection of 20 millimetres at a scale distance of 60 centimetres. The maximum safe R.M.S. current is 0·1 ampere.

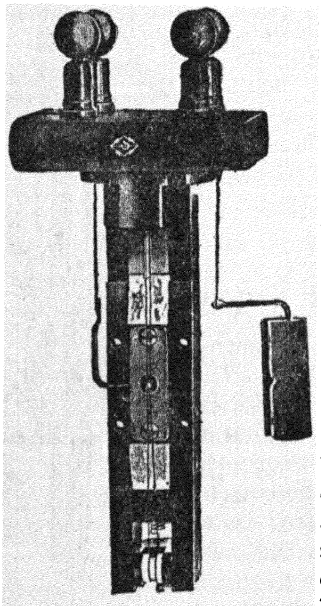


FIG. 225.

An electrostatic vibrator is shown in Fig. 225, and its circuit in Fig. 226. The magnetic field of the electromagnetic vibrator is replaced by an electrostatic field which is produced by applying the alternating pressure under test to the two metal plates F_1 and F_2. If the pressure exceeds about 2500 volts, the oscillograph is connected in series with a condenser C (or a system of condensers) which reduces the pressure to a suitable value at the oscillograph terminals, a, b. Two bronze strips, s_1, s_2, are stretched between the plates F_1, F_2, so that the strips are parallel to one another and to the plates. The strips are electrically insulated from one another by being joined at g and h to a silk thread t, which passes over an ivory pulley p; to this pulley is attached a spring q, which keeps the strips in a state of tension. The strips are charged with a definite potential by connecting the terminals e, j to a dry battery B mounted on an insulated stand. To make the potential of the strips definite in relation to the field plates, two condensers, C_1 and C_2, in series are connected to the points a, b across the field plates. The electrical centre k of the battery is connected to the point d between these condensers. If the

condensers C_1 and C_2 are equal, and the strips are midway between the plates, the amplitude of the vibrations of the strips will be equal when an alternating pressure is applied to the terminals of the oscillograph. A small mirror *m* which is bridged across the strips will therefore be deflected (since the strips, owing to their opposite electrification, will move in opposite directions in the field). This deflection is observed by focusing a beam of light on the mirror, the incident and reflected beams passing through a small opening w_1 in the plate F_1. An exactly similar opening, w_2, is cut in plate F_2 in order to maintain symmetry in the electric field. In practice it is found desirable to provide an electrical adjustment by making one of the condensers, namely C_2, variable. The electrostatic vibrator has an undamped period of about 0·00033 second, and the amplitude of a half wave vibration is about 6·2 millimetres per thousand volts at a scale distance of 60 centimetres with a D.C. potential of 400 volts between the vibrator strips. The electrostatic and electromagnetic vibrators are interchangeable.

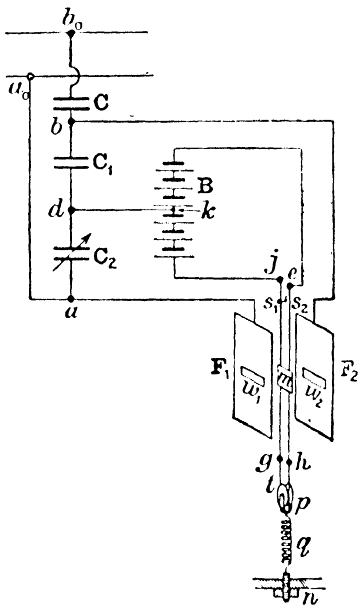


FIG. 226.

Calibration of Instruments measuring i^2.

Since the heating effect of a current is proportional to the square of the current, it is independent of its direction. So this heating effect may be used to measure either alternating or direct currents, and the alternating current measurements can be calibrated by means of direct current measurements.

There is, however, one point about such measurements which must be clearly understood. If the direct current

giving a particular deflection of any instrument measuring the square of the current be i, then the *maximum* value i_0 of the alternating current giving the same deflection is $i\sqrt{2}$.

This can be shown as follows :—

The instantaneous alternating current i at time t is, as we have seen,

$$i = i_0 \sin (pt + \varepsilon).$$

The heat produced in time dt by such a current flowing through a resistance R is

$$dH = \frac{R}{J} i^2 dt \text{ calories.}$$

If heat H is evolved in time T, we then have

$$H = \frac{R}{J}\int_0^T i^2 dt$$

$$= \frac{Ri_0^2}{J}\int_0^T \sin^2 (pt + \varepsilon)dt.$$

If T is taken as the time for one complete cycle $T = \frac{2\pi}{p}$ seconds.

So
$$H = \frac{Ri_0^2}{J}\int_0^{\frac{2\pi}{p}} \sin^2 (pt + \varepsilon)dt$$

$$= \frac{Ri_0^2}{J}\int_0^{\frac{2\pi}{p}} \frac{1 - \cos 2(pt + \ )}{2}dt$$

$$= \frac{Ri_0^2}{J}\Big[\frac{t}{2} - \frac{1}{4p} \sin 2(pt + \varepsilon)\Big]_0^{\frac{2\pi}{p}}.$$

When we take this definite integral, the two values of the sin term go out, since $\sin (2\pi + \varepsilon) = \sin \varepsilon$.

So putting again T for $\frac{2\pi}{p}$, we have

$$H = \tfrac{1}{2} . \frac{Ri_0^2 T}{J}.$$

But the heat evolved in time T by a direct current i is given by

$$H = \frac{Ri^2 T}{J}.$$

So if the steady current i produces in time T the same amount of heat as an alternating current whose maximum value is i_0, then

$$i^2 = \tfrac{1}{2} i_0^2, \text{ and hence } i = \frac{i_0}{\sqrt{2}}.$$

In this equation i is called the R.M.S. current (that is, The Root-Mean-Square current).

We could have expressed this fact more shortly, but without formal proof, by saying that the average value of $\sin^2\theta$ over a complete cycle is $\frac{1}{2}$.

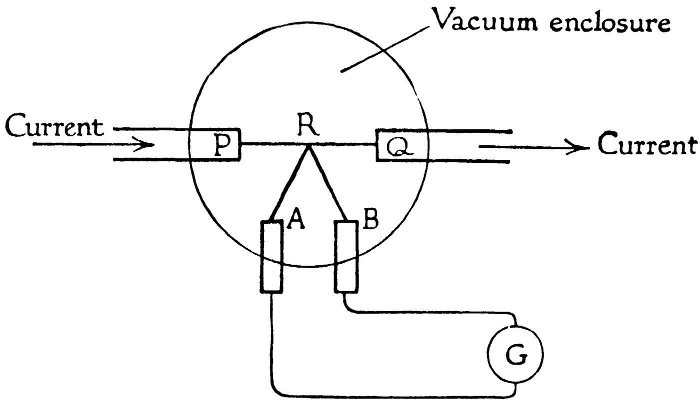


FIG. 227.

Thermo-ammeters

Instruments measuring currents or voltages by the heating effect are called thermo-ammeters or thermo-voltmeters. They are of two kinds. In the less accurate kind, such as the Cardew voltmeter, the actual expansion of a hot wire is magnified mechanically, and the resulting deflection, whose relation to i^2 is already known by independent calibration is observed by means of a pointer on a scale. The rate of heat production is proportional to i^2, but the temperature, and hence the deflection, depend on this in a complicated manner.

These instruments are apt to suffer from change of zero, and to have trouble with the mechanism for magnifying the expansion.

Another more satisfactory type uses what is called a vacuo-junction. This is simply a thermo-couple fixed to, but in-

sulated from, a wire carrying the current, as in Fig. 227, the whole instrument being *in vacuo* to eliminate air currents and improve the sensitivity. The thermo-couple is in series with a low-resistance galvanometer of high sensitivity, whose deflection is nearly proportional to the rate of production of heat, and can be calibrated against a steady current.

In the figure it is supposed that the alternating current to be measured flows through a wire PQ which has a thermo-couple AR–RB fixed near it at R. This system is *in vacuo*, and a sensitive low-resistance galvanometer is in series between A and B.

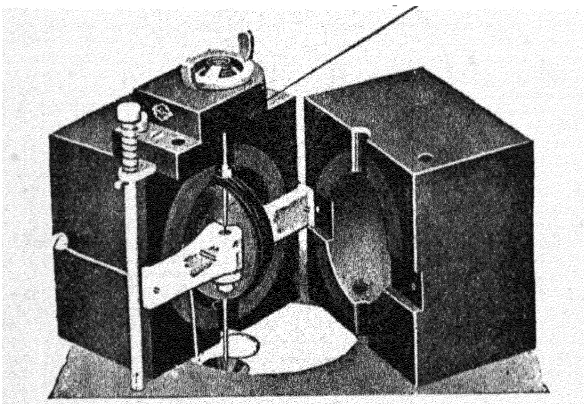


FIG. 228.

Dynamometers

The square of the current can also be measured by the force between two coils carrying it. Since the force between currents is proportional to $i_1 i_2$, the force between two coils both carrying the current i is proportional to i^2, and again if i be the steady direct current producing the same deflection as an alternating current whose maximum value is i_0, then

$$i_0 = \sqrt{2}i.$$

A typical dynamometer of this type is shown in Fig. 228.

Fig. 228 shows the moving coil, mounted on a single pivot, and free to turn between the two fixed coils which form an almost closed spherical chamber about the moving coil. The

method of pivoting ensures that the system is in equilibrium, and the design is such that levelling is unnecessary. The pivot is protected by being supported on a spring jewel, and each instrument is fitted with an external zero adjuster and coil lifting device. The controlling force is provided by a helical spring attached to the top of the moving coil, the current being led through this spring and a flexible ligament attached to the lower end of the coil. A damping vane is fitted to the moving system. The fixed coils are embedded in blocks of insulating material, and the clearances between fixed and moving coils are sufficiently small to give good air damping. Working parts and adjacent fittings, with the exception of a few small pieces of high-resistance material, are non-metallic, and eddy current errors are thus eliminated.

Electrostatic Voltmeters

For the measurement of alternating potentials without taking an appreciable current, the Quadrant Electrometer used idiostatically is the obvious instrument. Its use in this way has already been explained on p. 58. When it is set up to be used permanently for this purpose, it is usually called an Electrostatic Voltmeter.

Cathode-ray Oscillograph

The Cathode-ray Oscillograph measures high-frequency currents by a method employing the simplest principles directly. This method is thus far superior to all others in pure beauty.

The particles leaving the cathode in a highly exhausted discharge tube can be shown to be electrons. An electron, being a moving charge, may be regarded as a current, and experiences a force perpendicular to its motion when it moves through a magnetic field. The mass of an electron is about 9×10^{-28} gm., and its inertia is thus almost vanishingly small compared with that of the lightest vibration galvanometer or Duddell oscillograph. Its response to very high frequencies is thus incomparably better.

If the cathode rays are made to pass through a diaphragm and fall on a fluorescent screen of barium platinocyanide or

willemite, they produce a bright spot. If a coil carrying an alternating current is placed near the tube, the bright spot oscillates. If a moving film is made to receive the oscillations, the form of the current curve can be traced photographically on it. The particles can also be deflected by an electric field, and rapid alternating potentials, applied to metal plates on opposite sides of the stream of particles, can be measured when no current is flowing. The instrument can be used in this way for measuring sudden mechanical pressures with the help of piezo-electric crystals which are electrically polarized by pressure (see also p. 174).

Inductance in A.C. Circuits

If the current i in a circuit of inductance L is changing, then, as shown in Chap. IV, Part II, an E.M.F. of $L\frac{di}{dt}$ will be developed in the circuit. It acts always in such a direction as to oppose any change of existing conditions. If the current is increasing, the inductive E.M.F. opposes it. If the current is decreasing, the inductive E.M.F. supports it. Inductance thus behaves for a current in a manner analogous to the behaviour of mass for a moving body.

Let an alternating E.M.F. E given by

$$E = E_0 \cos pt$$

be applied to a circuit of negligible resistance. Let us consider in detail exactly what happens.

If there were no inductance as well as no resistance, the current would become infinite; but this tendency is checked at once if there is an inductance, since as soon as the current begins to grow a back E.M.F. appears and opposes its growth. It cannot stop the growth, however, for if the growth stops the back E.M.F. stops also, and the current increases again. It is evidently better to consider the equation which deals with the problem.

Things must automatically regulate themselves so that the back E.M.F. is equal and opposite all the time to the applied E.M.F., since the ohmic resistance is taken as negligible.

Thus
$$L\frac{di}{dt} = E_0 \cos pt$$

$$i = \frac{E_0}{L}\int_0^t \cos pt\,dt$$

$$= \frac{E_0}{pL} \sin pt$$

$$\therefore \quad i = \frac{E_0}{pL} \cos\left(pt - \frac{\pi}{2}\right) \quad . \quad . \quad . \quad . \quad (5)$$

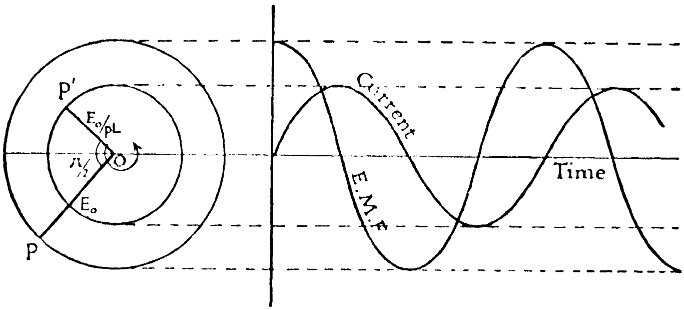


FIG. 229.

This equation tells us two things :—

1. The inductance L makes the maximum value of the current i_0 equal to $\frac{E_0}{pL}$. It thus behaves rather like a resistance of pL.

2. The current lags in phase by $\frac{\pi}{2}$ behind the E.M.F., *i.e.* the maximum value of the current is reached $\frac{1}{4}$ of a whole period after the maximum of the E.M.F.

The behaviour of a current through an inductance of negligible resistance is shown in Fig. 229.

It would be a mistake to call pL the *resistance* of L, since it only gives the ratio of the maximum value of the E.M.F. to the maximum value of the current, and it does not give the ratio of the instantaneous values. For this reason it has a different name, and is called the *Reactance*.

It will be shown later that when an inductance is used to cut down the current by means of its reactance it differs in

one most important particular from an ohmic resistance. *It wastes no energy*, for it produces no heat. It therefore gives the most economical way to regulate A.C. current.

It is important to notice that the reactance depends not only on the inductance but also on the frequency. The higher the frequency for a given inductance, the higher the reactance, and the smaller the current for a given E.M.F.

A small inductance can thus be used to restrict the flow of high-frequency currents without appreciably reducing low-frequency currents. Such an inductance is called a *high-frequency choke-coil.*

Capacity in A.C. Circuits

An alternating current behaves as if it can flow through a condenser, although a condenser must necessarily introduce a

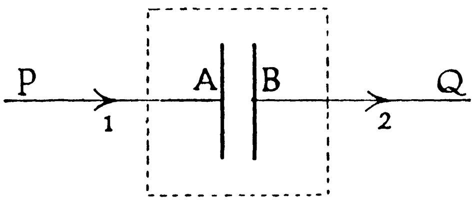


FIG. 230.

complete break into the circuit. It is important to understand how this happens.

In Fig. 230 imagine that the condenser-plates are contained in a closed box.

Let us now connect a battery to PQ, the outside terminals, so that P is positive.

Let us imagine that we can at any moment observe what is happening in the wires outside the box, but can see nothing of what is happening inside.

First positive charge flows along PA, causing a current in the direction of the arrow 1. The plate A is then charged positively. A negative charge is immediately "induced" or attracted on to the plate B, and this negative charge must flow in the direction QB. This is equivalent to a flow of positive charge out of B in the direction of the arrow 2. From the point of view of the observer, therefore, a positive

current flows into A and out of B, and thus appears to go through the condenser.

If we have simply connected a battery across PQ, this state of affairs cannot last. The charge on the plates will very quickly settle down to such a value that the difference of potential between A and B is the same as that of the battery, so that no current will flow. If, however, we replace the battery by a source of alternating potential it is clear that the original process described will be repeated as often as the voltage alternates. Thus unless any other factor (such as inductance) enters into the problem, the currents on both sides of the condenser will alternate with the potential difference of the source, and from the point of view of the observer the alternating current will flow through the condenser, since it will be flowing into one side while it is flowing out of the other.

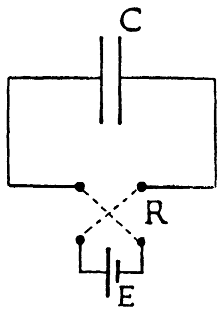


FIG. 231.

Reactance of Condenser

Let us consider now a special case in which a cell of E.M.F. E is used to charge a condenser of capacity C through a reversing switch R, as in Fig. 231.

Let the E.M.F. of the cell be applied to the condenser n times per second in each direction. When the switch is first turned on a charge CE flows into the condenser. When the switch is reversed this charge flows back, and a charge —CE flows on to the condenser. Thus the charge circulating in this time, $\frac{1}{2n}$ seconds, is 2CE. The average current is thus $\frac{2CE}{\frac{1}{2n}} = 4nCE$. The "effective resistance" of the condenser to an E.M.F. which goes through a cycle in this way n times per second is $\frac{E}{\text{current}}$, which

$$= \frac{E}{4nCE}$$

$$= \frac{1}{4nC}.$$

The condenser thus in this case behaves as if it were a resistance $\dfrac{1}{4nC}$ to the alternating current.

It can be shown that the similar "effective resistance" (which we may now call "Reactance") of the condenser to a sinusoidal alternating current of frequency n oscillations per second is $\dfrac{1}{2\pi nC}$, which is not very unlike the result we have found in our specially simple case.

In the simple case the graph showing E.M.F. against time would be as in Fig. 232, and the corresponding graph of current

FIG. 232.

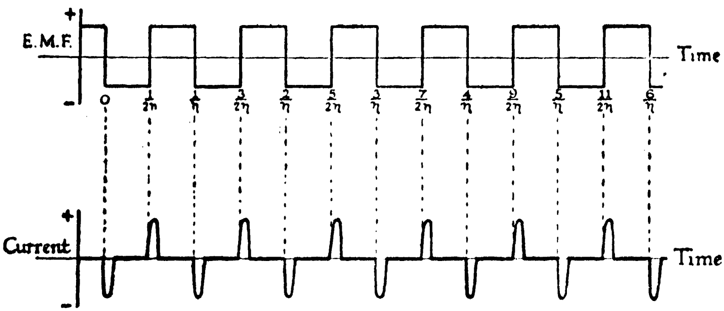


FIG. 233.

against time would be as in Fig. 233, immediately below. It is important to notice that the maximum value of the current always comes immediately after each change over, and not in the middle of the steady ranges of E. The current is thus again out of step with the E.M.F., being in this case ahead, instead of behind as with the inductance. This condition can now be treated mathematically, instead of in the artificial manner of this paragraph.

Exact Treatment of Condenser Problem

Let an E.M.F. E given by

$$E = E_0 \cos (pt + \epsilon)$$

be applied to a condenser of capacity C.

The charge Q on the condenser at time t is given by

$$\begin{aligned} Q &= EC \\ &= E_0C \cos (pt + \varepsilon). \end{aligned}$$

But i, the current, is equal to $\frac{dQ}{dt}$, by definition,

$$\therefore \quad i = -E_0pC \sin (pt + \varepsilon)$$

$$i = \frac{E_0}{1/pC} \cos\left(pt + \varepsilon + \frac{\pi}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Let us compare this equation with equation (5) which we got for the inductance L.

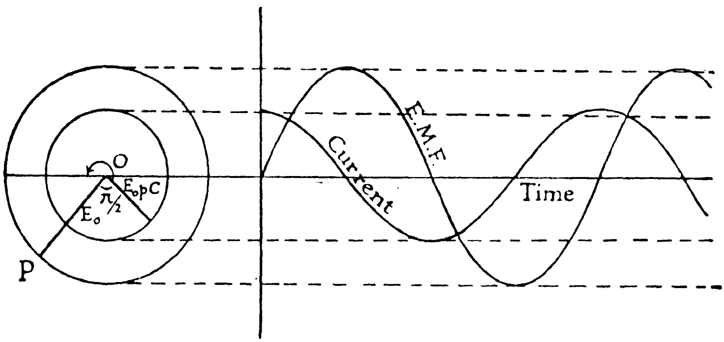


FIG. 234.

1. In place of the reactance pL, we have a reactance of $\frac{1}{pC}$ for the condenser.

2. In place of a retarding of phase of $\frac{\pi}{2}$ of current with respect to E.M.F., we have an advance of $\frac{\pi}{2}$.

3. The initial phase ε, which we have put in this time, has not affected the situation.

The effect of a condenser on an A.C. E.M.F. in a circuit without resistance or inductance is to allow a current to pass which is proportional to the frequency and the capacity, and to advance the phase of the current $\frac{\pi}{2}$ in front of the E.M.F.

The curves showing the relation of current to E.M.F. for a condenser are given in Fig. 234. The reactances of inductances and condensers are distinguished by being known as inductive reactances, and capacitative reactances, respectively.

Inductance with Resistance in A.C. Circuits

Real circuits always contain resistance.

In a circuit with resistance R and inductance L, we have at any time

(Back E.M.F.) + (ohmic potential drop round circuit) = (applied E.M.F.)

$$L\frac{di}{dt} + Ri = E_0 \cos pt \qquad \text{(7)}$$

In the capacity problem we added the initial phase ε, but it did not in any way affect the problem. So in future problems we can leave it out unless it is needed.

The equation (7) is a differential equation. The general methods of solving such equations are outside the scope of this book. We can, however, solve this one (*i.e.* make it into an ordinary equation connecting i and t) by a trick. Let us suppose that perhaps

$$i = A \cos (pt - \alpha)$$

will do it, if only the right values can be found for the constants A and α.

Then—

$$L\frac{di}{dt} + Ri \equiv -LAp \sin (pt - \alpha) + RA \cos (pt - \alpha)$$

$$\equiv A[-Lp \sin pt \cos \alpha + Lp \cos pt \sin \alpha + R \cos pt \cos \alpha + R \sin pt \sin \alpha]$$

$$\equiv A[(Lp \sin \alpha + R \cos \alpha) \cos pt - (Lp \cos \alpha - R \sin \alpha) \sin pt].$$

If this is to be identically equal to $E_0 \cos pt$, it must follow that the coefficient of $\cos pt$ must be E_0, and that of $\sin pt$ must be zero.

$$\therefore \left.\begin{aligned} A(Lp \sin \alpha + R \cos \alpha) &= E_0 \\ A(Lp \cos \alpha - R \sin \alpha) &= O. \end{aligned}\right\}$$

Squaring and adding, we get rid of the terms in $\sin \alpha$ and $\cos \alpha$, since $\sin^2 \alpha + \cos^2 \alpha = 1$, and the terms in $\sin \alpha \cos \alpha$ cancel each other.

Then $$A^2(L^2p^2 + R^2) = E_0^2.$$

$$\therefore \quad A = \frac{E_0}{\sqrt{R^2 + L^2p^2}}$$

and $$Lp \cos \alpha - R \sin \alpha = 0$$

so $$\tan \alpha = \frac{Lp}{R}.$$

The equation is thus

$$i = \frac{E_0}{\sqrt{R^2 + L^2p^2}} \cos (pt - \alpha) \quad . \quad . \quad . \quad (8)$$

Comparing this result with those for L and C alone, we see that the quantity $\sqrt{R^2 + L^2p^2}$ is analogous to the reactance, but that it contains resistance as well as inductance.

It is called the *Impedance*, Z.

We also see that the current lags by a phase-angle of $\tan^{-1}\left(\frac{Lp}{R}\right)$ behind the E.M.F.

We can say of this circuit that it has

Resistance R
(Inductive) Reactance Lp
Impedance $\sqrt{R^2 + L^2p^2}$.

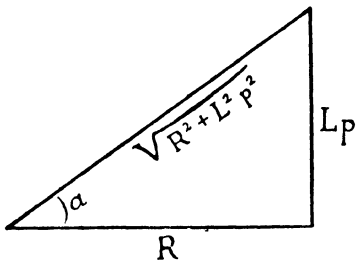


FIG. 235.

The effect of Lp, the reactance, on the phase-lag α is shown in Fig. 235.

Fig. 235 is *not* a vector diagram. Vector diagrams are considered below.

Capacity with Resistance in A.C. Circuits

By similar methods we can show that in a circuit containing resistance R and capacity C, but no inductance, the impedance is given by

$$Z = \sqrt{R^2 + \frac{1}{C^2p^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

and that the current is in advance of the E.M.F. by an angle α where

$$\tan \alpha = \frac{1}{CpR} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

In such a circuit we have

$$\text{Resistance } \mathrm{R}$$

$$\text{(Capacitative) Reactance } \frac{1}{\mathrm{C}p}$$

$$\text{Impedance } \sqrt{\mathrm{R}^2 + \frac{1}{\mathrm{C}^2p^2}}.$$

Capacity, Inductance, and Resistance combined in an A.C. Circuit

By similar methods again we can show that in a circuit containing capacity C, inductance L, and resistance R the impedance is

$$\mathrm{Z} = \sqrt{\mathrm{R}^2 + \left(\mathrm{L}p - \frac{1}{\mathrm{C}p}\right)^2} \quad . \quad . \quad . \quad (11)$$

and the phase-lag α is given by

$$\tan \alpha = \frac{\mathrm{L}p - \frac{1}{\mathrm{C}p}}{\mathrm{R}} \quad . \quad . \quad . \quad . \quad . \quad (12)$$

In such a circuit

$$\text{Resistance} = \mathrm{R}$$

$$\text{Reactance} = \left(\mathrm{L}p - \frac{1}{\mathrm{C}p}\right)$$

$$\text{Impedance} = \sqrt{\mathrm{R}^2 + \left(\mathrm{L}p - \frac{1}{\mathrm{C}p}\right)^2}.$$

Vector Diagrams

The behaviour of this circuit can be simply represented by rotating vectors placed with respect to each other as shown in Fig. 236.

In this figure OA represents the ohmic E.M.F., which is in phase with the current in the circuit.

OB represents the capacity E.M.F., which is $\frac{\pi}{2}$ behind.

OC represents the inductive E.M.F., which is $\frac{\pi}{2}$ in advance.

The resultant E.M.F. will be that corresponding to a vector which is the resultant of the first three.

The construction of this vector is shown in Fig. 237.

EOD is the phase-angle α, which in the case shown is a lag, since capacity is here (as shown in Fig. 236) more effective than inductance. Inductance and capacity vectors, being in the same straight line but in opposite senses, add algebraically, but of course subtract numerically. Their sum ED can, of course, be added to OD (OA in Fig. 236) by the parallelogram law.

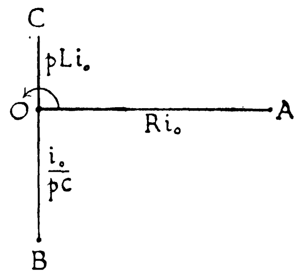


FIG. 236.

This method of computation, once it is understood, gives the results for a circuit very quickly and easily. In Fig. 237, OE thus represents the resultant E.M.F.

It should be noted that in the earlier Sections it was shown that the current in an Inductance lagged behind the E.M.F., whereas in a Capacity the current was in advance of the E.M.F. Thus in the earlier sections we started with the applied E.M.F. and considered the relationship of the current to it. This method seems to be easier when one first approaches the subject. In this section, however, we started with the current through the whole circuit, and considered the relationship of the various E.M.F.'s with the E.M.F. across R, which is in phase with the current. This method of drawing vector diagrams for the E.M.F.'s is usually in practice the best way of representing the circuit vectorially.

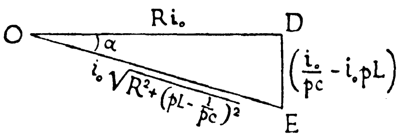


FIG. 237.

Power in A.C. Circuits

In a circuit containing inductance and resistance, if

$$E = E_0 \cos pt$$

then

$$i = \frac{E_0}{\sqrt{R^2 + L^2p^2}} \cos (pt - \alpha) \text{ where } \tan \alpha = \frac{Lp}{R}.$$

For shortness we may write this

$$i = i_0 \cos (pt - \alpha).$$

The work done by the current in an element of time dt is $\mathrm{E}idt$.

Hence if W = the rate of working, in ergs per second,

$$\mathrm{W} = n \int_0^{\frac{1}{n}} \mathrm{E}idt$$

$$= n\mathrm{E}_0 i_0 \int_0^{\frac{1}{n}} \cos pt \cos (pt - \alpha)dt.$$

If for shortness we put $pt = \theta = 2\pi nt$

$$\mathrm{W} = n\mathrm{E}_0 i_0 \int_0^{2\pi} \cos \theta \cos (\theta - \alpha) \frac{d\theta}{2\pi n}$$

$$= \frac{\mathrm{E}_0 i_0}{2\pi} \int_0^{2\pi} (\cos^2 \theta \cos \alpha + \cos \theta \sin \theta \sin \alpha) d\theta$$

$$= \frac{\mathrm{E}_0 i_0}{2\pi} \int_0^{2\pi} \left(\frac{\cos \alpha}{2} + \frac{\cos \alpha \cos 2\theta}{2} + \frac{\sin \alpha \sin 2\theta}{2}\right) d\theta$$

$$= \frac{\mathrm{E}_0 i_0}{2\pi} \left[\theta . \frac{\cos \alpha}{2} + \frac{\cos \alpha \sin 2\theta}{4} - \frac{\sin \alpha \cos 2\theta}{4}\right]_0^{2\pi}$$

$$= \frac{\mathrm{E}_0 i_0}{2\pi} . (\pi \cos \alpha)$$

$$= \left(\frac{\mathrm{E}_0}{\sqrt{2}}\right)\left(\frac{i_0}{\sqrt{2}}\right) \cos \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

Wattless Currents

Equation (9) is a most remarkable result. One would naturally have expected that the rate of working W would have been the product of the average current and the average E.M.F. But it is not. *The Phase-angle also appears.* Thus the greater the lag the less the power for given E.M.F. and current, since $\cos \alpha$ approaches 0 as α approaches $\frac{\pi}{2}$.

In particular, in a circuit in which R is negligible compared with L, so that the phase-lag is nearly $\frac{\pi}{2}$, or quarter of a period, *no power is used.*

It is for this reason that an inductance or choke-coil absorbs no power in an alternating current circuit. It should therefore be employed to regulate alternating current rather than a rheostat.

The part of the current which uses no power is known as the "wattless" component, and a current through a choke with negligible resistance is entirely wattless.

Similarly, a current through a circuit with capacity and no resistance is wattless, because again the current is $\frac{1}{4}$ period out of step; but this fact has little practical importance by itself, since all circuits have resistance.

General Properties of A.C. Circuits

The essential things to grasp about A.C. circuits are that capacity and inductance act in opposite ways, and can be used to compensate one another; and that increase of frequency increases the reactive effect of an inductance, but decreases the reactive effect of a capacity.

Neither capacity nor inductance can absorb and waste energy, as ohmic resistance wastes it, but either alone can prevent the energy from being transmitted. But for any given frequency and inductance value there is a particular capacity which will make the total reactance zero. In fact, for given values of any two of these three properties of the circuit—inductance, capacity, and frequency—there is a particular value of the third one which will cancel out the whole reactance.

Telegraph wires and electric cables have capacity. Submarine cables in particular have very large capacities of the order of 2 microfarads per mile. As we can see from (9) and (10), p. 453, this capacity has two effects. It increases the impedance and produces a phase-advance of the current. Both these effects are more marked as the frequencies get higher. When speech, which uses many frequencies, is being transmitted, they produce bad distortion. We can see from Questions 6 and 7 below that both effects can be cut out by the

inclusion of an inductance of the right value if only one frequency is used, and can be minimized if many frequencies are used.

In general it can now be seen that inductance and capacity have compensating effects. Capacity advances the phase of the current, and inductance retards it; the right combination of the two may keep the current in phase. Though either inductance or capacity alone would increase the impedance of the circuit, the combination of the two which prevents change of phase will also make the reactance zero. The impedance will then be equal to R.

Inductance is introduced in telephone lines by putting loading coils, of the right inductance and negligible resistance, at intervals along the line. It is introduced in cables by winding the cable in the form of a flat metallic tape round a central insulating core in the form of a solenoid of gigantic length, so as to distribute the inductance uniformly along it.

In the case of motors supplying power, the circuit has a large inductance but little capacity. This may produce a large lag and make most of the current wattless, thus reducing the power-factor of the motor. This can be compensated by putting a large bank of condensers across the dynamo, thus advancing the phase of the current and reducing the wattless component of the current. Only the in-phase component of the current does work. The wattless component does no work, but it has to flow along the mains, and it consumes power by heating the resistance. So the wattless component should be reduced as much as possible.

The Series Resonance Circuit

In a circuit containing inductance, capacity, and resistance, the impedance $\sqrt{R^2 + \left(Lp - \frac{1}{Cp}\right)^2}$ is smallest, being equal to R, when $p = \frac{1}{\sqrt{LC}}$, for in this case $Lp = \frac{1}{Cp}$, and the reactance is zero.

E.M.F.'s of this angular frequency will thus produce greater currents in the circuit than those of any other frequency, and the circuit is thus said to be *in resonance* to this frequency.

The circuit, Fig. 238, is known as a *series resonance circuit.* The absolute frequency, in oscillations per second, of the alternating current to which the circuit resonates, is thus

$$n = \frac{1}{2\pi\sqrt{LC}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

This formula should be compared with the formula for the frequency of a simple pendulum of length l under the action of gravity, which is

$$\frac{1}{2\pi\sqrt{l/g}}.$$

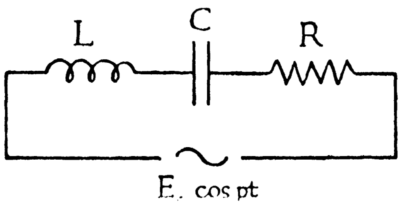


FIG. 238.

If there is no permanent applied alternating E.M.F., but a charge is suddenly given to the condenser, one of two things will happen.

If the resistance R is above a certain critical value, which can be shown to be $2\sqrt{\frac{L}{C}}$, the condenser will simply discharge without oscillations, the energy disappearing as ohmic heat.

If R is below this value, oscillations will persist, and the instantaneous current i at time t will be given approximately, if R is very small, by

$$i = i_0 e^{-\frac{Rt}{2L}} \sin\left(\frac{t}{\sqrt{LC}}\right).$$

The absolute frequency of these oscillations will, of course, be $\frac{1}{2\pi\sqrt{LC}}$ as before.

The series resonance circuit contains the essentials of the simplest form of wireless receiving circuit. The circuit resonates to the induced E.M.F. set up in the aerial by the incoming signal.

Numerical Examples

Question 1. R.M.S. voltage and current (see p. 352).

The effective current and voltage in an A.C. circuit are usually known as R.M.S. (Root Mean Square) for obvious reasons.

If the R.M.S. values of A.C. through a condenser are 200 volts and 10 amps. respectively, what are their maximum values?

Since $$E_{R.M.S.} = \frac{1}{\sqrt{2}}E_{max}$$

and $$i_{R.M.S.} = \frac{1}{\sqrt{2}}i_{max}$$

$$E_{max} = 200\sqrt{2} = 282{\cdot}8 \text{ volts}$$
$$i_{max} = 10\sqrt{2} = 14{\cdot}14 \text{ amps.}$$

Question 2. An alternating E.M.F. of R.M.S. value 100 volts and frequency 50 cycles is applied to a condenser of 2 microfarads capacity, the resistance of the circuit being negligible. What is the R.M.S. current? The impedance Z of the condenser is given by

$$Z = \frac{1}{2\pi nC}$$
$$= \frac{1}{2\pi \times 50 \times 2 \times 10^{-6}}$$
$$= \frac{5000}{\pi} \text{ ohms.}$$
$$i = \frac{\text{R.M.S. E}}{\text{I}}$$
$$= \frac{100}{5000/\pi}$$
$$= \frac{\pi}{50} \text{ amperes} = 6{\cdot}28 \times 10^{-2} \text{ ampere.}$$

Question 3. The alternating E.M.F. of Question 2 is applied to an inductance of 5 henries, in a circuit of negligible resistance. What R.M.S. current flows?

The impedance Z is given by

$$Z = 2\pi nL$$
$$= 2\pi \times 50 \times 5$$
$$= 500\pi$$
$$i = \frac{100}{500\pi} \text{ amperes} = 6{\cdot}37 \times 10^{-2} \text{ ampere.}$$

Question 4. The alternating E.M.F. of Question 2 is applied to a circuit containing the condenser of Question 3 and a resistance

of 1000 ohms. What R.M.S. current flows? The impedance Z of the circuit is given by

$$Z = \sqrt{\frac{1}{(2\pi nC)^2} + R^2}$$
$$= \sqrt{\left(\frac{5000}{\pi}\right)^2 + (1000)^2}$$
$$= 1000\sqrt{\frac{25}{\pi^2} + 1}$$
$$= 1{\cdot}88 \times 10^3 \text{ ohms.}$$
$$\therefore\ i = \frac{100}{1880} \text{ amperes} = 5{\cdot}32 \times 10^{-2} \text{ ampere.}$$

Question 5. The alternating E.M.F. of Question 2 is applied to a circuit containing the inductance of Question 3 and the resistance of Question 4. What R.M.S. current flows? The impedance Z of the circuit is given by

$$Z = \sqrt{(2\pi nL)^2 + R^2}$$
$$= \sqrt{(500\pi)^2 + 10^6}$$
$$= 1{\cdot}862 \times 10^3 \text{ ohms}$$
$$i = \frac{100}{1862} = 5{\cdot}37 \times 10^{-2} \text{ ampere.}$$

Question 6. The alternating E.M.F. of Question 2 is applied to a circuit containing the inductance of Question 3 and the condenser and resistance of Question 4. What R.M.S. current flows?

The impedance Z of the circuit is given by

$$Z = \sqrt{\left(\frac{1}{2\pi nC} - 2\pi nL\right)^2 + R^2}$$
$$= \sqrt{\left(\frac{5000}{\pi} - 500\pi\right)^2 + R^2}$$
$$= 1{\cdot}0002 \times 10^3 \text{ ohms}$$
$$i = \frac{100}{1000{\cdot}2} = 9{\cdot}98 \times 10^{-2} \text{ ampere.}$$

[Note that this is almost the same current that would have flowed through the resistance of 1000 ohms alone. The capacity and inductance have almost counteracted each other.]

Question 7. Express the lag or advance in Questions 2, 3, 4, 5, 6, of current with respect to E.M.F. in terms of phase-angle and of absolute time in seconds.

In Question 2 the current is $\frac{1}{4}$ period in advance, since the circuit is a pure capacity.

$$\therefore \text{ Phase-advance} = \frac{\pi}{2}$$

$$\text{Time-advance} = \frac{1}{4} \times \frac{1}{50} = 5 \times 10^{-3} \text{ second.}$$

In Question 3 the current is $\frac{1}{4}$ period behind, since the circuit is a pure inductance.

$$\therefore \text{ Phase-lag} = \frac{\pi}{2}$$

$$\text{Time-lag} = \frac{1}{200}'' = 5 \times 10^{-3} \text{ second.}$$

In Question 4 the phase-advance α is given by

$$\begin{aligned}
\tan \alpha &= \frac{1}{2\pi n CR} \\
&= \frac{1}{2\pi \times 50 \times 2 \times 10^{-6} \times 1000} \\
&= \frac{5}{\pi} = 1{\cdot}591 \\
\alpha &= 57^\circ\ 51' \\
&= \frac{57^\circ\ 51'}{180^\circ} \times \pi \text{ radians} \\
&= 1{\cdot}01 \text{ radians.}
\end{aligned}$$

The time-advance is $\frac{1}{50}''$ for 2π radians.

Here, then, time-advance

$$\begin{aligned}
&= \frac{1{\cdot}01}{2\pi} \times \frac{1}{50} \text{ second} \\
&= 3{\cdot}214 \times 10^{-3} \text{ second advance.}
\end{aligned}$$

In Question 5 the phase-lag is given by

$$\begin{aligned}
\tan \alpha &= \frac{2\pi n L}{R} \\
&= \frac{500\pi}{1000} = \frac{\pi}{2} = 1{\cdot}571 \\
&= 57^\circ\ 31' \\
&= 1{\cdot}004 \text{ radians.}
\end{aligned}$$

$$\begin{aligned}
\therefore \text{ Time-lag} &= \frac{1{\cdot}004}{2\pi} \times \frac{1}{50} \text{ seconds} \\
&= 3{\cdot}196 \times 10^{-3} \text{ second lag.}
\end{aligned}$$

In Question 6, the phase-lag is given by

$$\tan \alpha = \frac{2\pi nL - \dfrac{1}{2\pi nC}}{R}$$

$$= \frac{500\pi - \dfrac{5000}{\pi}}{1000} = -0{\cdot}02.$$

A negative lag is an advance.

$$\alpha = 0{\cdot}02 \text{ radian.}$$

[Note that it is only because the angle is small that radian measure = the tangent.]

$$\text{Time-advance} = \frac{0{\cdot}02}{2\pi} \times \frac{1}{50}''$$
$$= 6{\cdot}4 \times 10^{-4} \text{ second advance.}$$

Question 8. A telegraph wire has a capacity to earth of 0·015 microfarad per mile and is 200 miles long. It carries oscillatory currents of frequency 5000. What loading inductance is required altogether to make the A.C. resistance a minimum?

$$2\pi nL = \frac{1}{2\pi nC}$$

$$L = \frac{1}{4\pi^2 n^2 C}$$

$$= \frac{1}{4\pi^2 \times (5000)^2 \times 200 \times 0{\cdot}015 \times 10^{-6}}$$

$$= \frac{1}{300\pi^2}$$

$$= 3{\cdot}38 \times 10^{-4} \text{ henry}$$
$$= 338 \text{ microhenries.}$$

Question 9. What is the natural frequency of a circuit containing an inductance of 50 microhenries and a capacity of 0·0005 microfarad? What wave-length will it respond to?

$$n = \frac{1}{2\pi\sqrt{LC}}$$

$$= \frac{1}{2\pi\sqrt{50 \times 10^{-6} \times 0{\cdot}0005 \times 10^{-6}}}$$

$$= \frac{10^7}{\pi\sqrt{10}}$$

$$= 1{\cdot}007 \times 10^6.$$

It will respond to a wave-length of about $\frac{3 \times 10^{10}}{10^6}$ cm., or 300 metres.

Question 10. What would be the effect of (*a*) doubling, and (*b*) halving, the frequency in Question 6? Find also the effect on the phase-lag or advance.

(*a*)
$$\begin{aligned} I &= \sqrt{\left(\frac{2500}{\pi} - 1000\pi\right)^2 + (1000)^2} \\ &= 2554 \text{ ohms} \\ i &= \frac{100}{2554} = 3{\cdot}916 \times 10^{-2} \text{ ampere.} \end{aligned}$$

(*b*)
$$\begin{aligned} I &= \sqrt{\left(\frac{10{,}000}{\pi} - 250\pi\right)^2 + R} \\ &= 2598 \text{ ohms} \\ i &= \frac{100}{2598} = 3{\cdot}849 \times 10^{-2} \text{ ampere.} \end{aligned}$$

The original form of Question 6 gave a phase-advance of 1° 9′.

In case (*a*), the effect of inductance is increased, and the lag is given by

$$\begin{aligned} \tan \alpha &= \frac{1000\pi - \frac{2500}{\pi}}{1000} \\ &= \pi - \frac{2{\cdot}5}{\pi} \\ &= 2{\cdot}347 \\ \alpha &= 66° \, 55'. \end{aligned}$$

In case (*b*), the effect of capacity is increased, and the advance is given by

$$\begin{aligned} \tan \alpha &= \frac{\frac{10{,}000}{\pi} - 250\pi}{1000} \\ &= 2{\cdot}398 \\ \alpha &= 67° \, 22'. \end{aligned}$$

The enormous effect of changing the frequency should be noticed.

CHAPTER IX

UNITS AND DIMENSIONS

Dimensions—Dimensional Equations—Solution of Problems by Dimensional Equations—Quantities with Identical Dimensions not necessarily Identical—Dimensions of Mechanical Units—Dimensions of Electrical Units—Ratios of Electromagnetic and Electrostatic Units—Practical Units—Table of Electrical Dimensions and Units—Notes on the Use of Dimensions Table.

Dimensions

No length, however great, can have area; no area, volume. Length, area, and volume have different natures. No equation could have areas on one side and volumes on the other. The quantities expressed on the two sides of an equation must always be of the same nature.

On the other hand, the product of two lengths cannot be another length, but must be an area. The product of three lengths is a volume, and the product of a length and an area is a volume.

The nature of a quantity is determined by what are called its " Dimensions." We thus say that a volume has the dimensions of length cubed, and make this statement mathematically in what is called a " Dimensional Equation." If V stands for " volume," and L for " length," the dimensional equation relating volume and length is written

$$[V] = [L]^3.$$

Dimensional Equations

Quantities in dimensional equations are always surrounded by square brackets (which are tiring to write).

Thus Density, being $\frac{\text{Mass}}{\text{Volume}}$ is expressed

$$[D] = [M]\,[L]^{-3}.$$

A dimensional equation is written with negative indices rather than fractions—$[L]^{-3}$ rather than $\frac{1}{[L]^3}$.

Since velocity is $\frac{\text{Length}}{\text{Time}}$

$$[V] = [L]\,[T]^{-1}.$$

Since acceleration is $\frac{\text{Velocity}}{\text{Time}}$

$$[A] = [L]\,[T]^{-2}.$$

Since by Newton's Second Law of Motion, Force is rate of change of momentum, or $\frac{\text{Mass} \times \text{Velocity}}{\text{Time}}$

$$[F] = [M]\,[L]\,[T]^{-2}.$$

From which we see at once that we may also express this law by saying,

$$\text{Force} = \text{Mass} \times \text{Acceleration}$$

Kinetic Energy is Mass $\times$ (Velocity)^2, or $[M]\,[L]^2\,[T]^{-2}$. Work is Force $\times$ Length, or $[M]\,[L]\,[T]^{-2} \times [L]$, which again is $[M]\,[L]^2\,[T]^{-2}$, and we thus see that Work and Energy have the same nature.

If we divide a quantity by another of the same dimensions, we get a pure number, which our dimensional equation shows us to have zero dimensions. Pure numbers are thus the only quantities which are completely independent of our units of measurement. If a certain distance is three times as great as another distance, the ratio is 3 whether the distances are measured in miles, inches, centimetres, or light-years. True ratios are thus the only quantities which are ultimate in the sense that they do not depend upon our minds.

We have up to now used three kinds of dimension—Length, Mass, Time. These serve our ordinary purposes, though we have to regard temperature as having dimensions for certain problems. Heat, which we measure in calories, may, of course, be regarded as being energy.

The thermal capacity of a body is the heat required to raise

its temperature through 1° C. It is expressed in calories per degree, and must thus have the dimensions of $\dfrac{\text{Energy}}{\text{Temperature}}$.

Solution of Problems by Dimensional Equations

Dimensional equations, it is surprising to find, can be used to solve problems directly in certain circumstances.

Let us suppose that we are convinced for some reason that the time of swing of a simple pendulum depends upon its length, and the acceleration g of a falling body towards the earth, and that it is independent of the mass of the bob and of the length of the arc of swing. We know then that the quantities l and g alone occur in the equation for the time of swing, but we do not know to what power they are raised, or what constant numerical multiplier is needed.

We can, however, say,

$$\tau = \mathrm{K} \,.\, l^x g^y$$

where K is some constant, and x and y are the unknown powers to which l and g are raised.

Now, since K has zero dimensions and g is an acceleration, the dimensional equation corresponding to our equation for τ must be written

$$[\mathrm{T}] = [\mathrm{L}]^x \{[\mathrm{L}]\,[\mathrm{T}]^{-2}\}^y$$
$$\therefore\ [\mathrm{T}] = [\mathrm{L}]^{x+y}\,[\mathrm{T}]^{-2y}$$

For this equation to be true, the dimensions of [L] must be zero, and those of [T] must be 1.

So we have $\quad x + y = 0$
$$-\,2y = 1.$$

It thus follows that $\quad x = \frac{1}{2}$
$$y = -\,\tfrac{1}{2}.$$

So $\quad \tau = \mathrm{K} l^{\frac{1}{2}} g^{-\frac{1}{2}}$
$$= \mathrm{K}\sqrt{\frac{l}{g}}.$$

Thus we have arrived by a short cut at a well-known result, except that we cannot by this method find K.

Much more striking examples of this method can be given.

A drop of liquid under the action of its own surface-tension is in equilibrium in the form of a sphere. If it is displaced from the spherical shape it tries to return, overshoots the mark

and thus oscillates. The length of the period of oscillation is obviously a complicated thing to work out directly, but it can be shown to depend only on the surface-tension, the volume, and the density.

Thus $$T = KS^x V^y D^z$$

Surface-tension is Force per unit length, [M] [L] $[T]^{-2}$ divided by [L]. Thus

$$[S] = [M]\,[T]^{-2}$$

So $$[T] = \{[M]\,[T]^{-2}\}^x \{[L]^3\}^y \{[M]\,[L]^{-3}\}^z$$
$$= [M]^{x+z}\,[L]^{3y-3z}\,[T]^{-2x}.$$

So $$\left.\begin{aligned} x + z &= 0 \\ 3y - 3z &= 0 \\ -2x &= 1. \end{aligned}\right\}$$

$\therefore$ $$x = -\tfrac{1}{2}$$
$$y = z = \tfrac{1}{2}.$$

Thus $$\tau = K\sqrt{\frac{VD}{S}}.$$

Other interesting problems can easily be worked out by this method. A few examples are given.

The rate of flow of water through a tube of Q c.c. per second depends only on the pressure-gradient P in dynes per sq. cm. per cm. length, the radius r, and the viscosity η in dynes per sq. cm. per unit velocity gradient. It can be shown to be given by

$$Q = K \,.\, \frac{Pr^4}{\eta}.$$

Small ripples due only to surface-tension S have a velocity v depending on s, the density ρ and the wave-length λ.

$$v = K\sqrt{\frac{S}{\lambda\rho}}.$$

Shallow-water gravitational waves have a velocity depending only on g, and the depth d.

$$v = K\sqrt{gd}.$$

If it be thought that the density might affect the velocity, ρ can be included in the dimensional equation. It will be found to have zero index.

Deep-water gravitational waves have a velocity depending only on g and the wave-length λ.

$$v = K\sqrt{g\lambda}.$$

A particularly interesting and surprising case is that of the mass of a stone which can be lifted by a flowing river. This is of great importance in geology in the theory of denudation. If it is assumed that the mass M depends only on v, the velocity of the river, g, and ρ the density, then

$$M = K \frac{\rho v^6}{g^3}$$

and the mass of the stone varies with the 6th power of the velocity of the river. It is for this reason that the denudation of a river is largely due to flood-periods.

All these equations may be simply found by the method of dimensions.

Quantities with Identical Dimensions not necessarily Identical

Though quantities of the same nature must have the same dimensions (and the difficulty about the drawing of a muchness in *Alice in Wonderland* was due to the Dormouse's failure to realize this), it does not follow that quantities having the same dimensions are of the same nature. The frequency of an oscillation is the number of times it happens per second. Its dimensions are thus $[T]^{-1}$. The velocity-gradient of a stream is a velocity divided by a length; again $[T]^{-1}$.

But a velocity-gradient is different from a frequency. The difficulty arises from a deficiency in the notation. The idea of length is not concerned with frequency, but it is concerned with velocity-gradient.

For a velocity-gradient we take a length along our river and divide by time to find our velocity, and then measure a length perpendicular to the direction of flow, observe how the velocity changes at points along this length, and divide the change of velocity by the length. We thus divide one length by another in a direction at right angles to it, and allow both to disappear from the equation. There is thus a serious gap in the simple method of notation for all quantities involving vectors.

Unfortunately there is no accepted sign for temperature in dimensional equations; and consequently we must look very carefully into equations involving temperature. Here again problems may arise even if we do look carefully, for thermal capacity and entropy, which are certainly quite different, have the dimensions of energy divided by temperature.

Dimensions of Mechanical Units

The following table gives the dimensions of many non-electrical units. The name is given as c.g.s. (centimetre-gramme-second) where there is no ordinary name.

Nature of Unit.	Name of Unit.	Dimensions of Mass, M.	Dimensions of Length, L.	Dimensions of Time, T.
Area	sq. cm.	o	L^2	o
Volume	c.c.	o	L^3	o
Density	gm./c.c.	M^1	L^{-3}	o
Velocity	cm./sec.	o	L^1	T^{-1}
Acceleration	cm./sec^2.	o	L^1	T^{-2}
Force	dyne	M^1	L^1	T^{-2}
Work or Energy	erg	M^1	L^2	T^{-2}
Pressure	dyne/cm^2.	M^1	L^{-1}	T^{-2}
Frequency Angular Velocity	rev./sec. radian/sec.	o	o	T^{-1}
Velocity Gradient	c.g.s.	o	o	T^{-1}
Pressure Gradient	c.g.s.	M^1	L^{-2}	T^{-2}
Surface Tension or Surface Energy	c.g.s.	M^1	o	T^{-2}
Viscosity	c.g.s.	M^1	L^{-1}	T^{-1}
Power or Rate of doing Work	erg/sec.	M^1	L^2	T^{-3}
Action or Energy × Time	c.g.s.	M^1	L^2	T^{-1}
Angular Acceleration	radian/sec^2.	o	o	T^{-2}
Momentum	c.g.s.	M^1	L^1	T^{-1}
Moment of Momentum	c.g.s.	M^1	L^2	T^{-1}
Angular Momentum	c.g.s.	M^1	L^2	T^{-1}
Moment of Inertia	c.g.s.	M^1	L^2	o
Torque	c.g.s.	M^1	L^2	T^{-2}
Strain	c.g.s.	o	o	o
Stress	c.g.s.	M^1	L^{-1}	T^{-2}
Elasticity-modulus	c.g.s.	M^1	L^{-1}	T^{-2}
Compressibility	c.g.s.	M^{-1}	L^1	T^2
Diffusion-coefficient	c.g.s.	o	L^2	T^{-1}

Dimensions of Electrical Units

The dimensions of electrical and magnetic quantities are rather oddly related to those of mechanical quantities.

There are two complete systems of Absolute Units, as we have seen, one starting from Electric Charge and one from Magnetic Pole. The dimensions of any particular quantity are different on the two systems. Two distinct methods of dealing with them are in vogue. Both must be understood, since at present the allegiance of text-books is fairly evenly divided between them.

In the method used by Clerk Maxwell in his classical *Treatise on Electricity and Magnetism*, μ is regarded as having zero dimensions on the electromagnetic system and $[L]^{-2}\,[T]^2$ on the electrostatic system, while k has dimensions zero on the electrostatic system and $[L]^{-2}\,[T]^2$ on the electromagnetic system. All other units contain only M, L, and T.

This method leads to the somewhat surprising conclusion, categorically stated in Vol. II of Maxwell's book, that the practical unit of resistance, the ohm, "is expressed in the electromagnetic system by a velocity of 10,000,000 metres per second." It also leads to the conclusions that capacity on the electrostatic system, and inductance on the electromagnetic system, are lengths, and that specific resistance on the electrostatic system is a time.

In the other method, which was adopted later, k and μ are both regarded as having unknown dimensions on both systems, the only thing known about them being that the product of their dimensions is $[L]^{-2}\,[T]^2$ on both systems (as of course it is on Maxwell's system also).

By this method every electrostatic unit may have dimensions including M, L, T, and k, and every electromagnetic unit may have dimensions including M, L, T, and μ.

This method is on the whole more convenient in practical problems because it tells one what to do with k and μ, but the tendency at present is to return to Maxwell's method, and to interpret it in a rather different way.

The only power of M which appears in any electrical quantity is $+\frac{1}{2}$, and L appears fractionally in most of them.

It does not seem possible to attach any physical meaning whatever to a fractional dimension, and it has been suggested that the unknown dimensions of k and μ might in some way

remove this difficulty; or at least that their appearance is to be regarded as a signal that the problem is as yet unsolved. Probably, however, there were many unsolved problems in the Ptolemaic system of astronomy depending on epicycles; but, without being solved, they ceased to exist when the Copernican system appeared.

We will first calculate a few dimensions of quantities on both absolute systems, in order to show the method of attack, and demonstrate in the process that in any case $[k\mu]$ must have dimensions $[L]^{-2}\,[T]^{2}$.

We will use the method employing k and μ, since it certainly makes the working-out rather easier.

1. Electric Charge (Electrostatic system).

Charges q_1 and q_2, r cm. apart in a medium of dielectric constant k, exert a force $\frac{Q_1Q_2}{kr^2}$ on each other.

$$\therefore\ \left[\frac{Q_1Q_2}{kr^2}\right] = [\text{Force}] = [M]\,[L]\,[T]^{-2}.$$

Since r is a length, this gives

$$[Q] = [k]^{\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{3}{2}}\,[T]^{-1}.$$

2. Electric Current (Electrostatic system).

$$[i] = \left[\frac{Q}{T}\right] = [Q]\,[T]^{-1}$$

$$\therefore\ [i] = [k]^{\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{3}{2}}\,[T]^{-2}.$$

3. Magnetic Pole (Electromagnetic system).

$$\left[\frac{m_1m_2}{\mu r^2}\right] = [\text{Force}] = [M]\,[L]\,[T]^{-2}$$

$$[m] = [\mu]^{\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{3}{2}}\,[T]^{-1}.$$

4. Magnetic Intensity or Field Strength (Electromagnetic system).

$$[mH] = [\text{Force}]$$

$$[H] = \frac{[M]\,[L]\,[T]^{-2}}{[\mu]^{\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{3}{2}}\,[T]^{-1}}$$

$$= [\mu]^{-\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{-\frac{1}{2}}\,[T]^{-1}.$$

5. Current (Electromagnetic system).

The Magnetic Intensity H at the centre of a circular coil carrying a current i is given by

$$H = \frac{2\pi i}{r}.$$

where r is the radius of the coil.

$$\therefore\ [i] = [H]\,[L]$$

$$= [\mu]^{-\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{1}{2}}\,[T]^{-1}$$

We have now obtained i on both systems. Equating its dimensions on the two systems (since it is the same quantity in fact on either system)—

$$[k]^{\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{3}{2}}\,[T]^{-2} = [\mu]^{-\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{1}{2}}\,[T]^{-1}$$

$$\therefore \qquad [k\mu] = [T]^2\,[L]^{-2}$$

or

$$\left[\frac{1}{\sqrt{k\mu}}\right] = [\text{velocity}].$$

If any electrical unit whatever is worked out on both systems this result follows.

Now experiment shows that the ratio of the electromagnetic unit of current or charge to the electrostatic unit of current or charge is 3×10^{10} very nearly.

The measured velocity of light *in vacuo* is also 3×10^{10} cm./sec. very nearly.

If we divide the electromagnetic dimensions of current by the electrostatic dimensions we get the ratio

$$[L]^{-1}\,[T]\left[\frac{1}{\sqrt{k\mu}}\right]$$

and it thus follows, since this ratio is experimentally 3×10^{10}, that the velocity $\dfrac{1}{\sqrt{k\mu}}$ must be 3×10^{10} cm./sec.

Maxwell's calculations showed that electromagnetic waves should in fact be propagated with a velocity $\dfrac{1}{\sqrt{k\mu}}$. Since this value agreed with the observed velocity of light there was at once a very strong presumption that light was composed of electromagnetic waves. Other evidence has since confirmed this conclusion.

Ratios of Electromagnetic and Electrostatic Units

The ratio of the size of the electromagnetic unit to the size of the electrostatic unit of any electrical quantity is now known as soon as their dimensions are known on both systems. It is clear from the case of current discussed above that the following simple rule will give this ratio.

The power of c (the velocity of light, 3×10^{10} cm./sec.) in the ratio $\dfrac{\text{e.m.u.}}{\text{e.s.u.}}$ is the same as the power of T obtained by

dividing the dimensions on the e.m. system by those on the e.s. system.

Only four values of this power are ever found to occur. They are 2, 1, − 1, − 2. Thus the ratio $\frac{\text{e.m.u.}}{\text{e.s.u.}}$ can only have one of the four values, c^2, c, $\frac{1}{c}$, $\frac{1}{c^2}$, or 9×10^{20}, 3×10^{10}, $\frac{1}{3} \times 10^{-10}$, $\frac{1}{9} \times 10^{-20}$.

Practical Units

In many cases the absolute units have most inconvenient values. The electrostatic unit of charge, for example, is $\frac{1}{3{,}000{,}000{,}000}$ of a coulomb, and the electromagnetic unit of capacity is more than a billion times as great as the capacity of the earth.

A system of practical units has therefore been founded on two quite arbitrary choices.

The unit of current, the ampere, is chosen as being $\frac{1}{10}$ of the absolute e.m. unit.

The unit of energy, the joule, is chosen as being 10^7 ergs (about $\frac{3}{4}$ of a foot-pound, and $\frac{1}{4{\cdot}2}$ of a calorie).

The other practical units can be deduced from these values.

Thus the unit of charge, 1 coulomb, the charge which passes when 1 ampere flows for 1 second, is also $\frac{1}{10}$ of the e.m. unit.

The volt, the unit of potential, potential difference, or E.M.F. is the E.M.F. in a circuit in which 1 joule of work is done when 1 coulomb circulates.

It is thus the E.M.F. when 10^7 ergs are done when 10^{-1} e.m. units of charge circulate, or where 10^8 ergs of work are done when 1 e.m. unit circulates. Thus 1 volt $= 10^8$ e.m. units.

Since the e.m. unit of E.M.F. $= \frac{1}{3 \times 10^{10}}$ e.s. units, the volt is $\frac{10^8}{3 \times 10^{10}}$, or $\frac{1}{300}$ e.s. unit.

Similarly, it can be shown that the ohm is 10^9 e.m. units.

It is well worth remembering that 1 e.s. unit of potential = 300 volts, that 1 volt = 10^8 e.m. units, and that 1 ohm = 10^9 e.m. units, since these ratios turn up very frequently in practical problems.

Many of the practical units still have inconvenient sizes for many problems, and special prefixes are used to give secondary units. Thus "milli—" means $\frac{1}{1000}$ [*e.g.* milli-volt, milliamp.]. "Micro" means $\frac{1}{10^6}$ [*e.g.* microvolt, microamp., microhm, microfarad]. "Micro-micro" means $\frac{1}{10^{12}}$ [*e.g.* micro-micro-farad]. "Mega" means 10^6 [*e.g.* megohm, megadyne].

The dimensions of many electrical units, and their ratios to each other and to the practical units, are given in the following table, which is compiled according to Maxwell's method.

Notes on the Use of Dimensions Table

In this table c.g.s (centimetre-gramme-second) is given as the name of the unit whenever it has no special name, and when one of the absolute units is habitually used, rather than a unit derived from the ampere-joule system, the last column gives simply e.m.u. or e.s.u.

If k and μ are regarded as having dimensions on both systems, these dimensions can be deduced from the table in a simple manner.

We have seen that the ratio of electromagnetic to electrostatic dimensions for any unit must always be some multiple of $[L]^{-1}$ $[T]$ $[k]^{-\frac{1}{2}}$ $[\mu]^{-\frac{1}{2}}$. From this we have seen that the ratio $\frac{\text{e.m.u.}}{\text{e.s.u.}}$ must always be the same power of c as the power of T which appears in the dimensional ratio $\frac{\text{e.m.u.}}{\text{e.s.u.}}$. Thus it follows that for the e.s. system the power of $(k^{\frac{1}{2}})$ must be the same as the power of c in the ratio $\frac{\text{e.m.u.}}{\text{e.s.u.}}$. For the e.m. system the power of $(\mu^{\frac{1}{2}})$ must be the same as the power of c in the ratio $\frac{\text{e.s.u.}}{\text{e.m.u.}}$. This is perhaps a hard idea to express

*Q

clearly in words. It may be alternately stated in the following manner.

The power of k in an e.s. unit is $-\frac{1}{2}$ the power of c in the ratio $\dfrac{\text{e.s.u.}}{\text{e.m.u.}}$.

The power of μ in an e.m. unit is $-\frac{1}{2}$ the power of c in the ratio $\dfrac{\text{e.m.u.}}{\text{e.s.u.}}$.

For example,

$$\frac{\text{e.m.u. of charge}}{\text{e.s.u. of charge}} = \text{c}.$$

Hence dimensions of charge on the e.m. system may be written $[\mu]^{-\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{1}{2}}$, and on the e.s. system they may be written $[k]^{\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{3}{2}}\,[T]^{-1}$.

Again—

$$\frac{\text{e.m.u. of resistance}}{\text{e.s.u. of resistance}} = \text{c}^{-2}.$$

Hence dimensions of resistence on the e.m. system are $[\mu]\,[L]\,[T]^{-1}$, and on the e.s. system they are $[k]^{-1}\,[L]^{-1}\,[T]$.

For any one unit, the power of k on the e.s. system must clearly be numerically equal but opposite in sign to the power of μ on the e.m. system.

Knowledge of dimensions may be very enlightening even in simple cases.

For example, our rule gives the dimensions of H and i on the electromagnetic system as $[\mu]^{-\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{-\frac{1}{2}}\,[T]^{-1}$ and $[\mu]^{-\frac{1}{2}}\,[M]^{\frac{1}{2}}\,[L]^{\frac{1}{2}}\,[T]^{-1}$ respectively.

It is commonly stated that if a wire carries a current i at right angles to a magnetic field H, the force on a length x of wire is Hix dynes.

In this case we have

$$\text{H}ix = \text{Force}.$$

But the result of multiplying together the dimensions of H, i, and a length is found to be $[\mu]^{-1}\,[M]\,[L]\,[T]^{-2}$, instead of $[M]\,[L]\,[T]^{-2}$ as it should be.

It thus follows that the equation Hix = Force is wrong, and should be replaced by μHix = Force or

$$\text{B}ix = \text{Force}.$$

In actual cases it does not as a rule matter which equation one uses, since forces on currents usually act in media in which

the departure of μ from unity is negligible. But if a current did experience a force inside a magnetized medium, the force would be Bix, not Hix.

Similarly the force between currents i_1 and i_2, 1 cm. apart in a medium of permeability μ, is not $2i_1i_2$ dynes per cm., but $2\mu i_1i_2$ dynes per cm., since the dimensions of i^2 contain μ^{-1}.

Table of Electrical Dimensions and Units

Nature of Unit.	Electrostatic Dimensions.			Symbol Used.	Electro-magnetic Dimensions.			Ratio $\frac{\text{e.m.u.}}{\text{e.s.u.}}$	Name of Practical Unit.	Ratio of Practical Unit to useful Absolute Unit
	M	L	T		M	L	T			
Charge	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	T^{-1}	q	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	0	C	coulomb	10^{-1} e.m.u.
Current	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	T^{-2}	i	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	T^{-1}	C	ampere	10^{-1} e.m.u.
Electric Potential	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	T^{-1}	V	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	T^{-2}	1/C	volt	10^8 e.m.u. $\frac{1}{300}$ e.s.u.
Electric Intensity	$M^{\frac{1}{2}}$	$L^{-\frac{1}{2}}$	T^{-1}	E	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	T^{-2}	1/C	volt/cm.	10^8 e.m.u. $\frac{1}{300}$ e.s.u.
Magnetic Intensity	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	T^{-2}	H	$M^{\frac{1}{2}}$	$L^{-\frac{1}{2}}$	T^{-1}	C	oersted	e.m.u.
Electric Induction or Displacement	$M^{\frac{1}{2}}$	$L^{-\frac{1}{2}}$	T^{-1}	D	$M^{\frac{1}{2}}$	$L^{-\frac{3}{2}}$	0	C	tube	e.s.u.
Magnetic Induction	$M^{\frac{1}{2}}$	$L^{-\frac{3}{2}}$	0	B	$M^{\frac{1}{2}}$	$L^{-\frac{1}{2}}$	T^{-1}	1/C	gauss	e.m.u.
Magnetic Pole	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	0	m	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	T^{-1}	1/C	weber	e.m.u.
Magnetic Moment	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	0	M	$M^{\frac{1}{2}}$	$L^{\frac{5}{2}}$	T^{-1}	1/C	c.g.s.	e.m.u.
Magnetic Potential	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	T^{-2}	V	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	T^{-1}	C	c.g.s.	e.m.u.
Capacity	0	L^1	0	C	0	L^{-1}	T^2	C^2	farad centimetre	9×10^{11} e.s.u. 1 e.s.u.
Inductance	0	L^{-1}	T^2	L	0	L^1	0	$1/C^2$	henry	10^9 e.m.u.
Resistance	0	L^{-1}	T^1	R	0	L^1	T^{-1}	$1/C^2$	ohm	10^9 e.m.u.
Specific Resistance	0	0	T^1	S	0	L^2	T^{-1}	$1/C^2$	ohm/cm. cube	10^9 e.m.u.
Magnetic Flux	$M^{\frac{1}{2}}$	$L^{\frac{1}{2}}$	0	N	$M^{\frac{1}{2}}$	$L^{\frac{3}{2}}$	T^{-1}	1/C	line	e.m.u.
Intensity of Magnetization	$M^{\frac{1}{2}}$	$L^{-\frac{3}{2}}$	0	I	$M^{\frac{1}{2}}$	$L^{-\frac{1}{2}}$	T^{-1}	1/C	c.g.s.	e.m.u.
Dielectric Constant	0	0	0	k	0	L^{-2}	T^2	C^2	c.g.s.	e.u.s.
Permeability	0	L^{-2}	T^2	μ	0	0	0	$1/C^2$	c.g.s.	e.m.u.

*

MATHEMATICAL APPENDIX

Gauss's Theorem—Cavendish's Proof of the Inverse Square Law—Cells in Series and Parallel. Maximum Current—Electromotive Force of a Primary Cell.

Gauss's Theorem

THE Theorem states that—

" If any closed surface be taken in an electric field, the total normal induction across the surface is 4π times the charge enclosed by the surface."

Proof.—The electric induction across any area *in vacuo*, taken so that the direction of the field is normal to the area, is the total number of lines of force crossing the area; so that it may be regarded as corresponding to the magnetic flux.

If the field is uniform the induction may be regarded as the product of the field E and the area A.

If the field is not uniform, but is everywhere normal to the surface, it follows that the total induction is

$$\int E dA.$$

If the field is not uniform, and not normal to the surface, but makes at any point an angle θ with the normal, then the component of the field normal to the surface is everywhere

$$E \cos \theta.$$

Hence the total normal induction N, taken over the surface, is given by

$$N = \int^{0} E \cos \theta dA \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Really our N is the result of the summing of two entirely separate integrations.

If the total charge Q inside the surface is producing the total normal induction N, then we are trying to prove that

$$N = 4\pi Q \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Actually we must suppose that Q is made up of a lot of elements

dQ spread about anywhere inside the surface. If dN be the part of N due to dQ, then we must show that

$$dN = 4\pi dQ \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

But the proof of this involves the integration we have already mentioned in equation (1), for we must integrate all over the area of the surface to find the normal induction due to dQ.

But since the direction of the resultant intensity due to all the dQ's put together is not the same as that to a single dQ, we must perform the first integration not for θ but for another angle θ' between the field due to dQ alone and the normal to the surface dA.

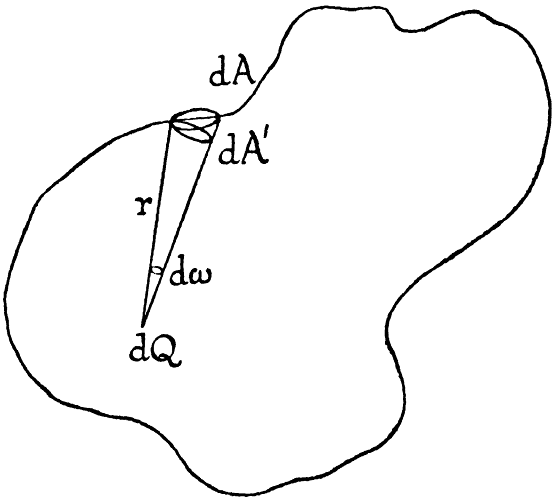


FIG. 239.

The useful form of equation (1) thus becomes

$$dN = \int^{0} dE \cos \theta' dA \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where dE is the intensity due to dQ alone.

Now suppose dA (Fig. 239) is distant r from dQ, and dA' is the area cut off perpendicular to r by straight lines joining dQ to the circumference of dA. Then $dA' = dA \cos \theta'$ and equation (4) becomes

$$dN = \int^{0} dE dA'.$$

But $dE = \frac{dQ}{r^2}$ by the inverse square law, so equation (4) becomes

$$dN = \int^0 \frac{dQ \, . \, dA'}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

But $\frac{dA'}{r^2} = d\omega$, the solid angle subtended by the area dA'.

The solid angle subtended by an area at a point is numerically equal to the area cut off on the surface of a sphere of unit radius surrounding the point by straight lines joining the point to the circumference of the area. So the solid angle subtended by a surface surrounding a point must be 4π, the area of the surface of a sphere of unit radius.

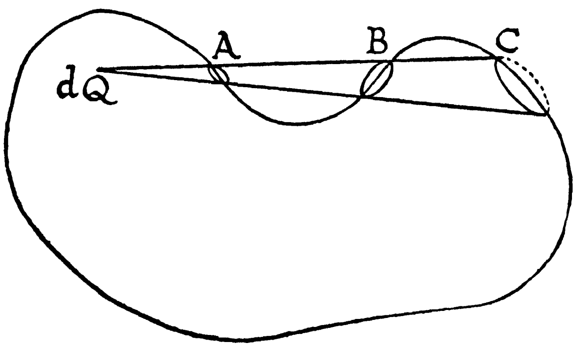


FIG. 240.

Hence substituting $d\omega$ for $\frac{dA'}{r^2}$ in (5) we get

$$\begin{aligned} dN &= \int^0 dQ d\omega \\ &= 4\pi dQ \\ \therefore N &= \int 4\pi dQ \\ N &= 4\pi Q \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6) \end{aligned}$$

which proves the theorem.

For a re-entrant surface the normal inductions in opposing surfaces (such as A and B in Fig. 240) will cancel out, so that only one surface, that at C, in the figure, will appear in the integration.

The fact that the charge on an electrostatic conductor is entirely on the surface may be shown to follow directly from equation (6).

For let our Gauss surface be taken entirely inside the conducting material of a conductor containing no insulated charged bodies. Then since the charges are at rest the intensity at every point inside the material is zero. Then the normal (and every other) component of the intensity is zero at every point of our Gauss surface.

Then the total normal induction across the surface is zero.

Then from equation (6) the total charge inside the surface is zero.

This goes on being true if we expand our Gauss surface, so long as none of it gets outside the material.

The moment it does get outside an intensity appears, and consequently the surface encloses some charge.

We may, therefore, say that:

" It follows from the Inverse Square Law that the charges on a conductor (containing no insulated charged bodies) are entirely on the surface."

The theorem has only been proved for conditions *in vacuo*, where k is 1.

In a medium of dielectric constant k the proof is valid if E is everywhere replaced by kE. This does not alter the result, but the k must reappear if we use the theorem to deduce the intensity outside a charged conductor, as is done in the section on the use of Gauss's Theorem in Chapter I, Part II.

Cavendish's Proof of the Inverse Square Law

" It follows from Cavendish's experiment that the law of force is that of the Inverse Square."

Proof.—It follows from our definition of the magnitude of a charge that the force between two charges Q and Q′, distant r apart, is given by

$$\text{Force} = QQ'f(r) \quad . \quad . \quad (1)$$

where $f(r)$ is some function of r which is as yet undetermined.

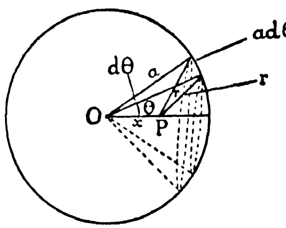


FIG. 241.

We require to prove that $f(r)$ is of the form $\dfrac{\text{constant}}{r^2}$.

Let us consider the potential V at a point P (Fig. 241) inside Cavendish's sphere, centre O, radius a. Let OP $= x$.

Consider a cone of axis OP, and semi-vertical angle θ.

It will cut the sphere in a circle, every point of whose circumference is the same distance, say r, from P.

Consider also another cone, same axis, with semi-vertical angle $(\theta + d\theta)$, cutting the sphere in a slightly larger circle, so that the distance on the circumference of the sphere, between the two circles, is $ad\theta$.

If the charge on the sphere is Q, uniformly distributed and all on the outer surface (as Cavendish's experiment shows), then since the charge per unit area is $\dfrac{Q}{4\pi a^2}$ and the area of the ring-element on the surface of the sphere cut off by our two cones is $(ad\theta) \times (2\pi a \sin\theta)$, it follows that the total charge on the ring-element, all of which is distant r from P, is

$$\frac{Q}{4\pi a^2} \times ad\theta \times 2\pi a \sin\theta$$

which simplifies to

$$\frac{Q}{2} \sin\theta d\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Now since the repulsive force between a charge Q and a unit charge r cm. away is $Qf(r)$, the work done in bringing the unit charge against the field from r to $(r + dr)$ is

$$- Qf(r)dr.$$

Hence the potential at a distance r from Q is

$$\int_\infty^r - Qf(r)dr$$

or, for simplicity,

$$\int_r^\infty Qf(r)dr \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

It therefore follows that the part dV of the potential at P due to the charge on the ring-element r away is [from (2) and (3)]

$$dV = \int_r^\infty \left[\frac{Q}{2} \sin\theta d\theta\right] f(r)dr.$$

But V is obtained by integrating over the whole surface, *i.e.* from $\theta = 0$ to $\theta = \pi$

$$\therefore V = \frac{Q}{2}\int_0^\pi \left[\int_r^\infty f(r)dr\right] \sin\theta d\theta \quad . \quad . \quad . \quad (4)$$

But from the triangle of sides a, r, x, we have

$$r^2 = a^2 + x^2 - 2ax \cos \theta.$$
$$\therefore\ 2rdr = +\ 2ax \sin \theta d\theta$$
$$\therefore\ \sin \theta d\theta = \frac{rdr}{ax}$$

and
$$\begin{cases} \text{when } \theta = 0,\ r = a - x \\ \text{when } \theta = \pi,\ r = a + x. \end{cases}$$

So we may get rid of θ in (4) by writing

$$V = \frac{Q}{2ax}\int_{(a-x)}^{(a+x)} \left[\int_r^\infty f(r)dr\right] rdr \quad . \quad . \quad . \quad . \quad (5)$$

Suppose now, for the sake of shortness and convenience, that instead of

$$\int \left[\int_r^\infty f(r)dr\right] rdr$$

we write $F(r)$,
then

$$\int_{(a-x)}^{(a+x)} \left[\int_r^\infty f(r)dr\right] rdr = F(a + x) - F(a - x).$$

Then we may rewrite (5) as

$$V = \frac{Q}{2ax}\left\{F(a + x) - F(a - x)\right\}.$$

Then

$$\frac{2a}{Q}Vx = F(a + x) - F(a - x).$$

Then, differentiating twice with respect to x, and treating V as a constant with respect to x, which we may do because the potential of the inner sphere of radius x is always the same as that of the outer sphere, since they have been in contact, whatever value of x we choose, and since V is due solely to the charge on the outside, there being none anywhere else,

$$0 = F''(a + x) - F''(a - x) \quad . \quad . \quad . \quad (6)$$

N.B.—This is the crucial step in this proof. It could not be made without the experimental evidence.

It follows from (6) that

$$F''(a + x) = F''(a - x).$$

That is $F''(r)$ is such a function that it is independent of the value of r; in other words, it is a constant.

$$\therefore \; F''(r) = A \quad \text{[Some constant.]}$$

We can integrate this twice to get back to $F(r)$.

$$\therefore \; F(r) = \frac{Ar^2}{2} + Br + C.$$

Substituting the real meaning of $F(r)$.

$$\therefore \; \int\left[\int_r^\infty f(r)dr\right]rdr = \frac{Ar^2}{2} + Br + C.$$

Differentiating

$$\left[\int_r^\infty f(r)dr\right]r = Ar + B.$$

$$\therefore \; \int_r^\infty f(r)dr = A + \frac{B}{r}.$$

Differentiating

$$f(r) = -\frac{B}{r^2}.$$

So that $f(r)$ is of the form $\dfrac{\text{constant}}{r^2}$

Q.E.D.

Since Cavendish's experiment shows that the charge on a spherical conductor was entirely on the surface, the second result, given the text of Chapter I, may be regarded as proved—namely :

" If the charge on a spherical conductor (containing no insulated charged bodies) is entirely on the surface, then the law of force is that of the inverse square."

Cells in Series and Parallel. Maximum Current

If we want to solve the general problem of how to arrange a given number of cells so as to get the biggest possible current through a given external resistance, we can choose between three ways of doing it; trial and error, guess-work, and differentiating. [A combination of all three is the most useful practically.]

Suppose n cells, each of E.M.F. E and internal resistance B, are arranged so as to have $\frac{n}{p}$ sets in series, each of p cells in parallel, so as to send a current i through an external resistance R.

Then the resistance of each set of p cells in parallel is $\frac{B}{p}$, so the resistances of $\frac{n}{p}$ of such sets in series is $\frac{nB}{p^2}$.

The total resistance in the circuit is $\left(R + \frac{nB}{p^2}\right)$.

The E.M.F. in the circuit is E times the number of paralleled sets (since these paralleled sets are arranged in series).

So the E.M.F. $= \frac{nE}{p}$.

So
$$i = \frac{\frac{nE}{p}}{R + \frac{nB}{p^2}}$$
$$= \frac{nE}{pR + \frac{nB}{p}}$$

To find the value of p giving a maximum value of i, we must differentiate i with respect to p, and find what value of p makes $\frac{di}{dp}$ vanish.

So
$$\frac{di}{dp} = nE\left[\frac{\left(R - \frac{nB}{p^2}\right)}{\left(pR + \frac{nB}{p^2}\right)^2}\right]$$

This is zero when
$$R - \frac{nB}{p^2} = 0$$

i.e. when
$$p = \sqrt{\frac{nB}{R}}.$$

This is the solution. But it is complicated in practice by the fact that p must be a whole number (since you cannot have $2\frac{1}{2}$ cells in parallel).

In an actual problem we have to try the integral values of p on either side of the calculated value, and choose the one giving the bigger current, if both divide exactly into n. If one divides into n and the other does not, the one which does is right.

If neither divides into n, the two nearest divisions on either side of the calculated value must be tried.

Example 7.—Suppose we have 144 cells, each of internal resistance 0·1 ohm and E.M.F. 2 volts. We want to send the biggest possible current through an external resistance of 0·2 ohm.

For maximum current

$$p = \sqrt{\frac{144 \times 0{\cdot}1}{0{\cdot}2}} = \sqrt{72} = 8{\cdot}485.$$

The calculated value of p lies between 8 and 9, both of which divide exactly into 144.

If $p = 8$,

$$i = \frac{nE}{pR + \frac{nB}{p}}$$

$$= \frac{144 \times 2}{8 \times 0{\cdot}2 + \frac{144 \times 0{\cdot}1}{8}}$$

$$= \frac{288}{1{\cdot}6 + 1{\cdot}8}$$

$$= 84\tfrac{12}{17} \text{ amps.}$$

If $p = 9$

$$i = \frac{144 \times 2}{9 \times 0{\cdot}2 + \frac{144 \times 0{\cdot}1}{9}}$$

$$= \frac{288}{1{\cdot}8 + 1{\cdot}6}$$

$$= 84\tfrac{12}{17} \text{ amps.}$$

So that in this case either value of p will give the same (maximum) value of i.

It will be found that any other value of p would give a smaller value of i.

Electromotive Force of a Primary Cell

Lord Kelvin suggested that the energy lost when metals go into solution in cells reappears as electrical energy. For example, the latent heat of formation of zinc sulphate is 37,730 calories per gramme-molecule. When a gramme-molecule of zinc sulphate is formed, 37,730 calories are *given up*.

The latent heat of formation of copper sulphate is —12,400 calorie per gramme-molecule. When a gramme-molecule of

copper sulphate is formed, 12,400 calories are *absorbed.* When it is decomposed they are given up.

When one faraday of charge, or 96,540 coulombs, passes, one gramme of hydrogen is liberated. As zinc is divalent, $\frac{1}{2}$ a gramme-molecule of zinc sulphate is formed, and $\frac{1}{2}$ a gramme-molecule of copper sulphate decomposed, when 96,540 coulombs pass.

Thus $\frac{37{,}730 + 12{,}400}{2}$ calories of energy should be liberated when 96,540 coulombs pass.

If all this heat goes into electrical energy, we have

$$1 \text{ coulomb does } \frac{25{,}065 \times 4{\cdot}182}{96{,}540} \text{ joules of work}$$

(taking J as 4·182 joules per calorie).

By the definition of the volt, the E.M.F. of the cell should then be this number of volts.

The value is 1·086 volts.

This value is very near the experimental value for the Daniell cell, but corresponding values are wrong for other cells.

Lord Kelvin's simple theory thus fails. It fails because it takes no account of the heat absorbed or evolved by the cell in working.

It may be worth while to state Lord Kelvin's equation which we have been using.

If H be the total heat lost in solution or formation per gramme-molecule of constituents that suffer chemical change in the cell, J be Joule's equivalent, n the valency, F the faraday, the charge that liberates one gramme of hydrogen, and E the E.M.F. of the cell

$$E = \frac{JH}{nF}.$$

Helmholtz altered this equation to bring in the heating effect.

It happens for the Daniell cell that $JH = nFe$, and it is for this reason that no thermal change is observed when it is working.

Imagine a reversible cell such as a Daniell, whose internal resistance is negligible so that when it passes a current we can neglect the loss of heat in the resistance of the circuit. [This assumption does not affect the argument.] Let its E.M.F. be E at absolute temperature T.

First let its temperature be raised from T to $(T + \delta T)$, so that its E.M.F. rises from E to $(E + \delta E)$. Let it generate a charge Q

at this temperature. It will produce electrical energy $Q(E + \delta E)$. Let it now cool to T, and pass Q coulombs in the opposite direction. The cell is now back in its original state. But it has given out $Q\delta E$ more energy than it has received. This energy can only have come from some heat other than Kelvin's heat.

By the second law of Thermodynamics a reversible machine (such as our cell) receiving heat h at a temperature $(T + \delta T)$ and rejecting it at temperature T can transform to work an amount of heat $\frac{h\delta T}{T}$.

If the cell absorbed heat H at temperature $(T + \delta T)$ we then have

$$\text{Energy obtained owing to } h = \frac{Jh\delta T}{T} = Q\delta E.$$

$$\therefore \; Jh = QT\frac{dE}{dT}.$$

Thus the energy per coulomb is $T\frac{dE}{dt}$ volts. So our final equation is

$$E = \frac{JH}{nF} + T\frac{dE}{dT}.$$

Note the distinction between H and h.

Thus if the E.M.F. of a cell rises with temperature, extra heat must come from somewhere, and the cell cools in order to provide this heat when it passes a current.

If the E.M.F. is diminished by rise of temperature the cell gives out heat when a current flows, and consequently its temperature rises. In the first case the E.M.F. is larger than Lord Kelvin's value; in the second smaller.

CHAPTER I.—ELECTROSTATICS

General Questions

1. Explain what is meant by electrical potential, and deduce an expression for the potential at an external point due to a charged sphere.

2. Deduce expressions for the potential, capacity and energy of a charged sphere in free space.

3. Deduce an expression for the energy stored in a parallel plate condenser by the following methods independently :

   (*a*) by considering the potential of the insulated plate;

   (*b*) by considering the work done in separating the plates against their mechanical attraction;

   (*c*) by considering the total energy stored in the dielectric.

4. Define potential and intensity, and show that the latter is the negative gradient of the former.

5. Find an expression for the capacity of a spherical shell condenser (*a*) *in vacuo*, (*b*) with a medium of dielectric constant *k*, whose thickness is half the difference between the radii of the spheres, placed between the shells.

6. Solve the problems of Question 5 for a long cylindrical shell condenser.

7. Deduce an expression for the mechanical force per unit area of surface of a charged conductor.

8. Explain the method of electrical images, and show by it how to find the surface density of charge at any point on an earthed conducting plane near an insulated point charge.

9. A point charge *in vacuo* is near an earthed sphere of different radius. Find an expression for the surface density of charge at any point on the surface of the sphere.

Numerical Questions

10. Assuming the earth to be a charged sphere in free space with a radius of $6{\cdot}4 \times 10^8$ cm. and an electric field of 3 volts per cm. at its surface, find the energy of its charge, and the heat that would be generated if it were completely discharged.

11. A parallel plate condenser has plates 2 cm. apart separated by alcohol of dielectric constant 30. The plates are charged to a p.d. of 30,000 volts. Find the force per unit area between them, and the energy stored in 1 c.c. of the alcohol.

12. A spherical shell condenser, with oil dielectric of dielectric constant 4, has radii of 10 cm. and 12 cm. for its shells. Find its capacity, and its energy when it is charged to a p.d. of 60,000 volts.

13. How will (*a*) the energy, (*b*) the force between the plates be affected when the difference between the plates of a parallel plate condenser is halved (1) with the charge on the insulated plate kept constant, (2) with the potential of the insulated plate kept constant?

14. Solve the problem of Question 12 for a cylindrical shell condenser 20 cm. long instead of a spherical shell condenser.

15. A horizontal earthed conducting plane has a point charge of 10 e.s.u. at a distance of 20 cm. from it. Find the surface density of charge, and the electric intensity, at a point 30 cm. from the foot of the perpendicular dropped from the point charge on to the plane.

16. Find the intensity and potential at a point 10 cm. vertically above the point where the surface density was found in Question 15.

17. A point charge of 10 e.s.u. is 30 cm. from the centre of an earthed conducting sphere of radius 10 cm. Find the surface density of charge (*a*) on the nearest part of the sphere's surface, (*b*) at a point at the end of a radius perpendicular to the line joining the point charge to the centre of the sphere.

18. Find the potential and intensity at a point midway between the charge and the centre of the sphere in Question 17.

19. The plates of a condenser of capacity 1 microfarad are joined by a resistance of 1,000,000 ohms. How long does the p.d. between the plates take to fall to half its original value?

20. Find the capacity of a condenser whose p.d. is reduced by 20% in 5 seconds by a resistance of 10^{12} ohms.

21. A cable contains a core of radius 1 cm., insulated by composition of dielectric constant 4 of thickness 2 cm., with an outer covering of lead. What is its capacity per kilometre?

22. Find the radius of a water-drop which will just remain suspended, with a charge of 1 electron, in the earth's field of 3 volts per cm.

23. Two drops, of radii 1 mm. and 2 mm. respectively, are at potentials of 15,000 volts and 7,500 volts respectively. What

is the change of energy in ergs if they coalesce, (*a*) neglecting surface tension, (*b*) assuming the surface tension to be 74 dynes per cm.?

24. An oil drop of density 0·8 gm./c.c. charged with 1 electron falls under gravity with a uniform velocity of 1 mm. per sec. Find the vertical field required to make it move upwards with the same velocity. The viscous force acting on a sphere of radius r moving with velocity v through a medium of coefficient of viscosity η is $6\pi r\eta v$. The density of air is 0·0013 gm./c.c. $g = 981$ cm./sec. Charge on electron $= 4\cdot774 \times 10^{-10}$ e.s.u. λ/η for air $= 1\cdot7 \times 10^{-4}$ c.g.s. unit.

CHAPTERS II AND III.—ELECTROMAGNETISM

General Questions

25. Explain very carefully the conception of the magnetic shell. What is meant by its strength? Find the magnetic potential due to a shell of strength σ at a point at which it subtends a solid angle ω.

26. Find an expression for the magnetic potential due to a shot bar-magnet at any point near it.

27. Explain carefully the foundations of Ampère's definition of unit current. Do you know of any equivalent definition?

28. Using only the fundamental definition of unit current, find an expression for the magnetic intensity at any point on the axis of a circular coil of wire carrying a current.

29. Two similar coils with their axes coincident and their planes parallel are separated by a distance equal to their radius r. Each coil has n turns. Show that the field at the centre of the system is $\dfrac{32\pi ni}{5\sqrt{5}r}$ oersted when i absolute units of current are flowing. (This arrangement is called the Helmholtz galvanometer.)

30. Find an expression for the magnetic intensity at any point on the axis of a short solenoid.

31. Show from the result of Question 30 that the intensity along the axis of a long solenoid with n turns per cm. carrying current i, is $4\pi ni$.

32. Show that the work done in taking a unit magnetic pole round a current i is $4\pi i$, and is independent of the path followed by the pole.

33. Explain what is meant by the line-integral of a magnetic field, and how it is determined in the case of a long straight wire and a short solenoid. What is meant by the line integral of an electric field?

34. Deduce an expression for the magnetic intensity at any point near a long straight wire carrying a current.

35. Find an expression for the mechanical couple on a flat coil of wire carrying a current in a magnetic field.

36. Find an expression for the force between a magnetic pole and a current element.

37. Find the force between two straight wires of length l lying parallel r cm. apart and carrying currents i and i_2 e.m.u.

38. An electron of charge e and mass m is projected with velocity v perpendicular to a magnetic field H. Show that it will describe a circle of radius $\frac{mv}{He}$ cm.

Numerical Questions

39. What current in amperes would give a deflection of 30° to a Helmholtz galvanometer of 20 turns to each coil of radius 10 cm. where H = 0·18 oersted? (See Question 29.)

40. A solenoid 10 cm. long of radius 1 cm. has 200 turns of wire carrying 1 ampere. Find the intensity at a point on the axis (a) in the centre of the coil, (b) in the plane of one end.

41. Find (a) the magnetic potential, (b) the magnetic intensity at a point on the axis of a single coil of wire of radius 5 cm. carrying a current of 20 amperes. The point is 10 cm. from the centre of the coil.

42. The coil of a sensitive galvanometer has 200 turns of average area 2 sq. cm. There is a radial magnetic field of 2000 gauss. A current of 1 microampere gives a deflection of 20°. Find the couple in dyne-cm. required to give 1 radian twist to the suspension.

43. The wires 0·5 cm. apart and a metre long carry currents of 5 amperes each. Find the force between them in grammes weight.

44. An electron is projected with a velocity of 10^9 cm./sec. at right angles to a magnetic field of 100 oersted. What will be the radius of the circle it describes?

Charge of electron $= 4{\cdot}774 \times 10^{-10}$ e.s.u.

Mass of electron $= 9 \times 10^{-28}$ gm.

45. The mass of a moving electron is $\dfrac{m}{\sqrt{1 - v^2/c^2}}$ when m is its rest-mass, v is its velocity and c is the velocity of light (3×10^{10} cm./sec.). What must its velocity be in order that it may describe a circle of radius one millimetre in a field of 15,180 gauss?

46. A current of 20 amperes flows along a straight wire. Find the magnetic intensity 10 cm. from the wire, and the work done in taking a unit pole round the wire.

47. A solenoid of radius 5 cm. contains 200 turns in each 10 cm. of its length, and carries 5 amperes. Find the flux through it, and the magneto-motive force in a 200 cm. length of it.

CHAPTER IV.—ELECTROMAGNETIC INDUCTION

General Questions

48. State the laws of electromagnetic induction, and show how Faraday's Law may be expressed as a differential equation.

49. Explain how electromotive force is related to change of magnetic flux, and magnetomotive force to change of electric flux.

50. Explain the principles and use of the earth inductor.

51. A coil rotates in a magnetic field, and a rectifying device makes the current unidirectional. This current is measured (*a*) by means of its heating effect and (*b*) by means of its electrolytic effect. Show that the ratio of the measured value of (*a*) to that of (*b*) is $\pi/2\sqrt{2}$.

52. A copper disc of radius r with its plane perpendicular to a magnetic field H rotates n times per second. Show that an E.M.F. of $\pi r^2 n$H e.m.u. acts between the axle and the circumference. (This arrangement is known as Faraday's disc.)

53. Explain carefully what is meant by self-induction and mutual induction.

54. Show that the energy stored in a circuit of self-inductance L in which a current i is flowing is $\frac{1}{2}Li^2$.

55. What determines the rate at which currents increase or decrease in a circuit to which an E.M.F. is applied? Find an expression for the current i at time t in a circuit of inductance L and resistance r to which an E.M.F. e is applied at time $t = 0$.

Numerical Questions

56. A wire is drawn in 0·2 second across the air-gap of an electromagnet of pole-strength 50,000 webers. What is the average E.M.F. induced?

57. A wire 100 cm. long, forming part of a circuit of resistance 2 ohms is moved at 200 cm./sec. across a field of 10,000 oersted. Find the E.M.F. in volts and the force opposing the motion of the conductor.

58. A wire 20 cm. long lies across two smooth rails between which there is a p.d. of 10 volts. The resistance of the wire is 0·1 ohm. A magnetic field of 2000 oersted acts at right angles to both wire and rails. What velocity will the wire finally attain?

59. In Question 58, the mass of the wire is 1 kilogram and the coefficient of friction between wire and rails is 0·25. What will now be the final velocity?

60. A disc like that in Question 52 has radius 5 cm. and revolves 100 times per second in a field of 2000 oersted. What E.M.F. is generated?

61. A coil has an inductance of 0·5 henry. Find the induced E.M.F. when the current grows at the rate of 10 amperes in $\frac{1}{100}$ of a second.

62. When the circuit is broken a current of 20 amperes collapses in 0·05 second, the induced E.M.F. being 100 volts. What is the inductance?

63. A p.d. of 1 volt is applied to a coil of resistance 1 ohm and inductance 1 henry. What is the current after one-tenth of a second? How long does it take the current to reach half its final value?

64. How much energy is stored in the coil of Question 63 when the current has reached its steady value?

65. A coil of resistance 2 ohms and inductance 0·1 henry is connected by a double-pole switch across mains of constant p.d. 20 volts. Just after connection is made, what is the induced E.M.F. and rate of change of current when the current is 5 amps.?

66. The switch in Question 65 is taken out, and instantly the coil is short-circuited through a negligible resistance. What is now the induced E.M.F. and rate of change of current when the current is 5 amps.?

67. A condenser of capacity 0·001 farad is being charged from 200-volt mains through a resistance of 5 ohms of inductance 0·1 henry. The rate of change of current is 200 amps. per second when the current is 10 amps. Find the p.d. between the plates

of the condenser at this instant, and at a later instant when the charge on the condenser has increased by 0·02 coulomb. Find also the induced E.M.F. when the current is 10 amps.

68. A circuit has resistance 1 ohm and negligible inductance. Its mutual inductance with another circuit, in which the rate of charge of current is 100 amps./sec., is 0·1 henry. What is the induced current?

69. An earth inductor has 200 turns of area 500,000 sq. cm. The resistance when it is in series with a flux-meter is 100 ohms. H is 0·18 oersted, and the dip is 67°.

What charge will pass when the coil is turned through 180° with its plane (*a*) vertical, (*b*) horizontal? (The maximum possible charge passes in each case.)

70. An earth inductor of area 1,000 sq. cm. gives a maximum throw of 2 microcoulombs when it is turned through 180° with its plane vertical, and 5 microcoulombs when its plane is horizontal. Find the dip. (Total resistance 100 ohms.)

71. If H is 0·2 oersted, how many turns has the earth inductor in Question 70?

72. A circular coil of mean radius 20 cm. with 200 turns makes 50 revs. per second about a diameter perpendicular to the lines of force in a magnetic field of 100 oersted. Find the maximum current and the average current (independent of sign). (The resistance of the coil is 4 ohms.)

73. An airship of 40 metres diameter has a coil of wire of 100 turns wound round it. It is moving at 120 kilometres per hour in a circle of radius 500 metres. H is 0·18 oersted. What is the maximum E.M.F. generated in the coil?

CHAPTER V.—MAGNETISM

General Questions

74. Distinguish between magnetic induction and intensity of magnetization, and establish an equation connecting them.

75. Distinguish between permeability and susceptibility, and establish an equation connecting them.

76. Draw a curve illustrating the relation between magnetic flux and magnetic induction when a piece of iron is taken through a cycle of magnetization. What information can be obtained from such a curve?

77. Distinguish between diamagnetism and paramagnetism, and give a short account of the theories which attempt to account for them.

78. Explain what is meant by ferromagnetism, and give a general account of a theory to explain it.

79. Explain carefully what is meant by magnetomotive force, and discuss its relation to electromotive force.

80. What is meant by the "magnetic circuit"? What are its resemblances to, and differences from, the magnetic circuit?

81. What is the effect of an air-gap on the magnetization of an iron ring by a field of given strength? How do you account for this effect?

82. Deduce an expression for the energy stored in unit volume of a magnetic field. For a field of given strength, is this energy greater in air or in iron?

83. Explain the meaning of the term "hysteresis" both in a general way and in connection with the magnetization of iron? Why is it specially important in the latter case?

84. Explain how you would obtain a hysteresis curve for a short cylindrical bar of iron.

85. Explain how you would obtain a hysteresis curve for an iron ring.

86. Show how to obtain an expression for the energy-loss in ergs per c.c. per cycle for a specimen of iron. Of what importance is this quantity?

87. Sketch the hysteresis curves you would expect to obtain for various ferromagnetic substances, and explain their significance.

Numerical Questions

88. A bar magnet is in the form of a cylinder of length 10 cm., radius 1 cm., and magnetic moment 10,000 units. Find its pole-strength and intensity of magnetization, and the number of lines of induction inside it.

89. An iron wire of 1 sq. mm. cross-section is placed in a solenoid 30 cm. long with 300 turns. The wire is 20 cm. long. A magnetometer on the axis of the solenoid, lying in a magnetic easterly direction, has its deflection increased from 0° to 30° when the wire is put in the solenoid. $H = 0{\cdot}18$ and the solenoid's current is 5 amperes. The magnetometer is 30 cm. from the middle point of the iron. Find the permeability of the iron.

90. An iron ring of circumferential length 60 cm. and cross-section 5 sq. cm. is wound with 1,700 turns of wire carrying

2 amperes, and the flux is found to be 200,000 lines. Find the permeability.

91. An iron ring 100 cm. long and 10 cm. in cross-section has a gap 1 mm. wide. How many turns of wire carrying 1 ampere (ampere turns) are needed to produce a field of 10,000 oersted in the gap? (μ = 1000.)

92. An iron ring, consisting of four quadrants whose cross-sections are 2, 3, 4 and 5 sq. cm. respectively, is 100 cm. long, and requires 100 turns carrying 2 amperes to produce a flux of 5,000 lines. Find the permeability of each section of the iron in these conditions.

93. An air-gap of 1 mm. width is made in the ring of Question 92. What current will be required to produce the same flux? The gap is in the quadrant of smallest cross-section.

CHAPTERS VI AND VII.—ELECTROMETERS AND GALVANOMETERS

General Questions

94. Describe the moving-coil galvanometer, and deduce an equation for its behaviour with steady currents showing that the deflection is proportional to the current passing.

95. What is meant by the sensitivity of a galvanometer? What factors determine it?

96. Explain what is meant by a "ballistic" galvanometer. Deduce an equation giving the relation between the charge sensitivity, the current sensitivity, and the time of swing.

97. What is meant by "critical" damping? How may it be obtained on a moving-coil galvanometer used for measuring steady currents?

98. What is the relation between damping and circuit resistance in a moving-coil ballistic galvanometer? How may the damping error of a ballistic galvanometer be compensated?

99. Describe a moving-magnet galvanometer. What are the advantages and disadvantages of this type when compared with the moving-coil galvanometer?

100. Describe the string galvanometer. For what purposes is it specially useful?

101. Describe the Grassot fluxmeter. Wherein does it differ from an ordinary moving-coil galvanometer?

102. Describe the construction of a quadrant electrometer, and deduce an expression for the deflection of the needle in terms of the potentials of the quadrants and that of the needle.

103. How may the capacity of an electrometer be determined experimentally?

104. Describe the gold-leaf electroscope. How may it be adapted to give quantitative results?

105. How would you compare the sensitivities of the quadrant electrometer and the gold-leaf electroscope to (*a*) charge, (*b*) potential? What results would you expect to find?

106. Why does the sensitivity of a quadrant electrometer to quadrant potential not increase indefinitely with needle potential, but reach a maximum and then decrease? How could you determine approximately this optimum value without plotting the whole curve?

107. How would you measure or compare high resistances with a quadrant electrometer?

108. How may capacities be compared (*a*) with a quadrant electrometer, (*b*) with a ballistic galvanometer? Which way do you prefer, and why?

109. How may dielectric constants be measured?

110. What is meant by the natural leak of an electrometer? What can be done to prevent it as far as possible from having ill effects?

111. Describe the tilted gold-leaf electroscope. What purpose is served by the tilt, and what is the instrument specially useful for?

112. Describe a method of measuring ionization currents.

113. Write an essay on the developments of the quadrant electrometer, paying special attention to the uses of particular types.

114. Describe the string electrometer. For what purposes is it specially useful?

Numerical Questions

115. A ballistic galvanometer has an undamped period of 12 seconds. Its deflection for a steady current of 1 microcoulomb is 250 mm. at 1 metre. What would be its throw for a charge of 1 microcoulomb?

116. The galvanometer of Question 115 has a resistance of 100 ohms. When a resistance of 900 ohms is in series with it, each swing is four-fifths of the former swing in the same direction. On open circuit each swing is 2% less than the former swing in the same direction. Find the damping coefficient, and the fractional diminution of successive swings in the same direction

when (*a*) the series resistance is 400 ohms, (*b*) the galvanometer is short-circuited.

117. A quadrant electrometer is in parallel with a guard-ring condenser whose plates have an area of 400 sq. cm. When the plates are 10 cm. apart the deflection is 250 mm. When the distance between the plates is reduced to 4 mm. the deflection is decreased to 100 mm. Find the capacity of the electrometer.

118. In Question 117, what would be the deflection when the plates were 1 cm. apart?

119. In Question 117, the introduction of a slab of dielectric 2 mm. thick between the plates reduces the deflection to 85 mm. What is the dielectric constant of the medium?

CHAPTER VIII.—ALTERNATING CURRENTS

General Questions

120. Explain what is meant by a sinusoidal current, and how it may be produced.

121. What are the advantages in the use of alternating current for the supply of electrical power over long distances?

122. What are the main types of instrument used for the measurement of alternating current? Discuss the principles on which each type depends.

123. In what sense may an alternating current be said to pass through a condenser? What general effect has an alteration in capacity of the condenser on the current passing for a given alternating E.M.F.?

124. What is the effect of inductance on a circuit to which an alternating E.M.F. is applied?

125. The effects of inductance and capacity on the alternating current in a circuit are sometimes said to oppose one another. Explain this.

126. Distinguish between reactance, resistance, and impedance, explaining carefully the meaning of each term.

127. Deduce expressions for the reactance of (*a*) a condenser, (*b*) an inductance.

128. Deduce an expression for the impedance of a circuit containing resistance, capacity, and inductance.

129. Using the result of Question 128, find the impedance of (*a*) a circuit containing capacity and resistance, and (*b*) a circuit containing inductance and resistance.

130. What is meant by electrical resonance? Deduce an expression for the resonance frequency of a circuit containing inductance and capacity.

131. How may (*a*) the capacity of a long-distance cable, (*b*) the inductance of a dynamo supplying A.C. be overcome?

132. What do you mean by phase-lag? What causes it and how may it be counteracted?

133. What determines the power consumed in an A.C. circuit?

134. What is the most economical method of controlling the quantity of A.C. current delivered by supply mains to an arc-lamp, and why has this method its special properties?

135. Explain carefully the term "wattless current."

136. Describe the 3-phase system of power-supply. What are its advantages?

Numerical Questions

137. What R.M.S. current is sent through a non-inductive resistance of 20 ohms by 200-volt (R.M.S.) 50-cycle A.C.?

138. If the circuit of Question 137 has an inductance of 0·03 henry, what current is sent and what is the phase-lag?

139. If the circuit of Question 137 has a series capacity of 300 microfarads, what current passes, and what is the phase-advance?

140. If the circuit of Question 137 has an inductance of 0·03 henry and a capacity of 300 microfarads in series with it, what current passes, and what is the phase difference between E.M.F. and current?

141. What would be the results of Questions 138, 139, and 140 if the A.C. were 100-cycle instead of 50-cycle?

142. At what rates would the circuits of Questions 138, 139, 140 be working?

143. At what rates would the circuits of Question 141 be working?

144. A submarine cable has a capacity of 0·4 microfarad per kilometre. If the frequency of transmission is 1,000 cycles, what should be the loading inductance per kilometre?

145. A dynamo supplying *n*-cycle alternating current has inductance L. What capacity should be in parallel with the supply to counteract phase-lag?

146. An arc of resistance 2 ohms requires 20 amperes at 60 volts, and is run off 100-volt 50-cycle A.C. mains. What inductance should be placed in series with the arc-lamp, and what will be the resulting phase-lag?

147. A circuit is required to resonate to a frequency of 193,000 cycles. Its capacity is 0·001 microfarad. What inductance is required? If the inductance takes the form of a coil of radius 5 cm., about how many turns will be needed?

148. The potential of each of the three wires of a 3-phase system has a maximum of 283 volts above or below earth. Find the maximum and R.M.S. value of the p.d. between any two wires.

149. A choke-coil takes 1 amp. from 200-volt 50-cycle mains. It has an iron core of length 100 cm. and radius 2·5 cm. μ may be taken as 2,000. Obtain an approximate value for the number of turns.

150. A circuit contains a condenser of reactance 60 ohms at 100 cycles. The mains are 200-volt 50-cycle. How much power is absorbed (*a*) if the condenser does not leak, (*b*) if it allows 0·5 ampere to pass at 200 volt D.C.?

ANSWERS

PART I

31. 12 e.s.u.
32. 62·64 e.s.u.
33. 24 dynes.
34. 83·92 e.s.u.
35. $1{\cdot}431\,\frac{Q}{a^2}$ along perpendicular bisector.
$1{\cdot}914\,\frac{Q^2}{a^2}$ dynes outwards along line of diagonal.
36. $0{\cdot}02854\frac{Q}{a^2}$ e.s. units parallel to the base.
37. 0·2 e.s.u. 2 e.s.u.
38. 37·5 dynes. 8·33 dynes. 13·57 ergs.
39. 48 volts per cm.
40. 11,310 volts per cm.
41. $2{\cdot}65 \times 10^{-13}$ coulomb per sq. cm.
42. $-1{\cdot}366 \times 10^{6}$ coulombs.
$-1{\cdot}92 \times 10^{9}$ volts.
43. $\frac{1}{3} \times 10^{-5}$ coulombs.
$2{\cdot}65 \times 10^{-9}$ coulomb/sq. cm.
44. 9×10^{5} cm.
45. 0·00375 mf.
46. 1·608 calories.
47. $\frac{3}{2}$.
48. 277778 ergs. $\frac{10}{13}$ mf.
49. 3306 volts.
50. 10 e.s.u
51. $-\frac{1}{8}$.
52. $1{\cdot}34 \times 10^{-7}$ calorie.
53. 0·278 ergs.
54. $2{\cdot}045 \times 10^{19}$ cm./sec.
$4{\cdot}423 \times 10^{-10}$ second.
55. 1000 volts. 0·002 coulomb.
56. 1190 volts.
57. 0·0811 gm.
58. $\frac{5}{8}$ cm.
59. 5 joules. 2·5 joules.
60. 408 ergs. 85·6 e.s.u

74. 8 dynes
$\frac{\sqrt{21}}{5}$ oersted for like poles.
$\frac{\sqrt{13}}{5}$ oersted for unlike poles.
75. (*a*) $\frac{8}{9}$ oersted.
(*b*) $\frac{2}{5\sqrt{5}}$ oersted.
(*c*) $\frac{1}{8}$ oersted
76. $1 : \sqrt{7} : \sqrt{\tfrac{7}{3}}$
77. 2465 units. 123·25 webers.
78. 68° 5′.
79. 66°; 50° 48′.
80. 2424 units. 0·18 oersted.
81. 15·12″.
82. (*a*) 788·6 dynes repulsion.
(*b*) 788·6 dynes attraction.

99. 110 volts.
100. 11·23 amps.
101. $\frac{1}{4}$ amp. 800 ohms.
102. 84%.
103. 0·24 volt. 4 volts.
104. 900 ohms.
105. 300 volts. $6\frac{2}{3}$%.
106. 190 volts. 186 volts. 94·2%.
107. 1·2 ohms.
108. $\frac{4}{33}$ ohm shunt.
109. $\frac{100}{999}$ ohm shunt.
110. 4·50 cm. 9·35 cm.
111. 9·00 volts.
112. 8·66 michroms/inch cube
113. 0·0711 cm.
114. 4·5 volts, 1 ohm.
115. 1·5 volt. 1·0 volt.

116. 2·25 amp.
117. 1 ohm. 0·5 amp.
118. $4\frac{160}{231}$%.
119. 28·3 ohms.
121. $56\frac{1}{4}$, 18, $5\frac{5}{8}$, $2\frac{29}{31}$, $1\frac{1}{2}$, $\frac{5}{16}$ amps.
122. $\frac{5}{8}$ ohm. $\frac{3}{8}$ and $\frac{1}{8}$ amp.
123. $1\frac{3}{8}$ ohm. $\frac{5}{14}$, $\frac{2}{7}$, $\frac{1}{14}$, $\frac{1}{7}$ amp.
124. 1 amp. $2\frac{20}{21}$ ohms.
125. (*a*) 1·8, 0·4, 2·2 amps.
(*b*) $\frac{6}{7}$, $-\frac{1}{14}$, $\frac{11}{14}$ amp.
(*c*) $\frac{14}{31}$, $-\frac{17}{62}$, $\frac{11}{62}$ amp.
126. (*a*) 1·0, 1·0, 2·0 amps
(*b*) $\frac{4}{7}$, $\frac{1}{7}$, $\frac{5}{7}$ amp.
(*c*) $\frac{12}{31}$, $-\frac{7}{31}$, $\frac{5}{31}$ amp
127. $-1\frac{5}{7}$ volt.
128. 13 in parallel.
129. 2 in series.

141. 22 minutes.
142. 22·7 ohms.
143. 4, $\frac{1}{4}$.
144. 0·1 ohm. 1 kilowatt.
145. $\frac{3}{8}$.
146. 0·546 ohm.
147. 33·3° C.
148. $800\sqrt{2}$ volts.
149. $2\sqrt[3]{2} : 1. = r_1/r_2$.
$1 : \sqrt[3]{2}. = l_1/l_2$.
150. 33·13 c.p.
151. $33\frac{1}{3}$ ohms.
152. 30 : 3 : 10 : 1.
153. $8·5 \times 10^{-4}$ calorie per sq. cm. per sec. per 1° C.

172. Back E.M.F. 1·6 volts. Resistance 8 ohms.
173. 0·615 gm.
174. + 0·055 volt.
175. 3·76 gm. 41·8 litres.
176. 33·64.
177. 2·05 volts.
178. 1·5 volts.
179. 3.
180. 200 volts, 100 watts.
181. 15
182. $\frac{9}{17}$, $\frac{3}{5}$ amp.
183. $8\frac{171}{178}$ volts. $1\frac{7}{178}$ in 7-ohm; $\frac{25}{89}$ in 6-ohm; $\frac{30}{89}$ in 5-ohm; $\frac{75}{178}$ in 4-ohm.

200. 0·503 oersted.
201. 7·555 cm. 0·0433.
202. 2865.
203. 0·0955.
204. 35° 16′.
205. 8000 dynes.
206. 360 dynes. 15° N. of E.
207. 80 dyne-cm.
208. 6283 lines, 20 oersted, 25 lines.
209. 3979 webers.
210. 1257 dynes.
211. 5·556 cm.
212. 0·4 dyne.
213. $4\pi \times 10^{-4}$ amp.
214. 5×10^4 cm. per sec.
215. $13\frac{1}{3}$ dynes.
216. $0·4\pi$ volt.
217. 10,000 volts.
218. $\frac{1}{100}$ coulomb.
219. 32 coulombs.
220. 0·5 volt.
221. 0·0054 volt.
222. 0·02 amp.
223. $1·6875 \times 10^{-3}$ volt.
224. 4×10^{10} dyne-cm.
225. 10·2 volts. 210·2 volts.
226. 16·2 volts. 194 volts. 1·5 ohms.
227. 197 volts.
228. 194 volts. $\frac{194}{197}$.
229. $\frac{3025}{3152}$.
230. $\frac{198}{197}$.
231. $4\frac{1}{2}$ ohms. 90 volts.
232. 230 volts. 250 volts.
233. 2667.
234. $83\frac{1}{3}$ amps., $101\frac{2}{3}$ volts.
100 amps., 100 volts.
$133\frac{1}{3}$ amps., $96\frac{2}{3}$ volts.

ANSWERS

PART II

10 $1\cdot311 \times 10^{22}$ ergs

11 $\frac{3}{32\pi} \times 10^5$ dynes; same number of ergs.

12. $4\cdot8 \times 10^6$ ergs.

13. halved: doubled: unaltered: quadrupled.

14 $4\cdot39 \times 10^6$ ergs.

15. $\sigma = 6\cdot79 \times 10^{-4}$ e.s.u
$E = 8\cdot54 \times 10^{-3}$ e.s.u.

16 $E = 9 \times 10^{-3}$ e.s.u.
$V = 8 \times 10^{-2}$ e.s.u.

17. $\frac{1}{80\pi}$ e.s.u. 0·002013 e.s.u

18 $\frac{8}{21}$ e.s.u $\frac{44}{2205}$ e.s u

19. $\log_e 2$

20. 20·17 e.s.u.

21. $1\cdot823 \times 10^5$ e.s.u.

22. $1\cdot051 \times 10^{-5}$ cm.

23. 52·9 ergs. 46·6 ergs

24 420 e.s.u $1\cdot26 \times 10^6$ volts per cm

39. 0·0578 amp.

40. 24·7 oersted. 12·5 oersted.

41. 1·32 e.m.u. 0·225 oersted.

42. 0·023 dyne-cm.

43. 0·102 gm. wt.

44. 0·579 cm.

45. 2×10^{10} cm./sec.

46. 0·4 oersted. 8π ergs.

47. $1000\pi^2$ lines. 8000π e.m.u

56 $\frac{n}{100}$ volt

57. 2 volts 10^5 dynes.

58. $2\cdot5 \times 10^4$ cm. per sec.

59. $9\cdot675 \times 10^3$ cm. per sec

60 $\frac{\pi}{20}$ volt.

61. 500 volts.

62. 0·25 henry

63. 0·095 amp. $\log_e 2$ second

64. 5×10^6 ergs.

65. 10 volts. 100 amps per sec.

66 Same as 65.

67. 130, 150, 20 volts

68 10 amps.

69 $3\cdot6 \times 10^{-3}$ coulomb
$8\cdot5 \times 10^{-3}$ coulomb

70. 68° 12′

71. 50.

72. 4π amps. $2\pi^2$ amps

73. $0\cdot048\pi$ volt.

88 $m = 1000$. $I = \frac{1000}{\pi}$. $N = 4\pi \times 1000$ e.m.u

89. 1109.

90. 562

91 1590.

92. 996, 664·5, 498, 398

93. 4 amps.

115 131 mm

116 0·1115 0·653 0·1285

117 $\frac{150}{\pi}$ cm

118. 160 cm.

119 4

137. 10 amps.

138. 9·05 amps. 25° 14′

139 8·83 amps. 27° 57′.

140. 9·99 amps. 3° 26′ advance

141 7·29, 9·67, 8·28 amps 43° 16′ lag, 14° 51′ advance, 34° 5′ lag.

142. 1037, 1560, 1989 watts

143. 1062, 1869, 1372 watts

144. 0·0633 henry

146. 53° 4′.

147 0·00068 henry. 37

148. 489, 346 volts

149. 360.

150 0, 100 watts.

PART III

CHAPTER I

ATOMIC PHYSICS, 1919–32

THE way to understand atomic physics is Rutherford's way, unless you are a really exceptional mathematician.

This way is to carry in your mind a very clear picture of a simple kind of sub-sub-microscopic billiards; to do for any one problem as much mathematics as you can get on the back of an envelope; to understand that little completely, and to enjoy yourself.

Rutherford had no patience at all with complicated mathematics or complicated experiments, or indeed with anything complicated. One could not take notes at his lectures, because everything he said seemed as simple as a cannon at billiards (and much of it was atomic billiards anyway). The danger was that unless one actually did the experiments he was talking about, or at least saw them, one could forget his words as a very vivid dream is sometimes forgotten by breakfast-time.

Where Rutherford's methods failed (where de Broglie, Heisenberg, Schroedinger, Dirac, and others have succeeded) he did not bother. There was always plenty more to get on with by the simple methods he had made his own.

He once said something like the following, when wave-mechanics began to hold the field: "It 's funny to be told, after working with α-particles for twenty years, that the α-particle is an illusion; but I don't mind so long as I am allowed to have my illusion concentrated in a very small space."

His playful and slightly satirical attitude to mathematics was well displayed in an incident I saw. He was giving one of those chatty lectures which great men gave from time to time at the Cavendish laboratory. The research staff and Part II students were in the habit of attending, and there were often good discussions. Suddenly he was struck by an idea, a bright but entirely unproved hypothesis. He turned suddenly to C. G. Darwin (later Sir Charles Darwin, head of the N.P.L.) who was sitting at the end of the front row, and said: "You ought to be able to prove that, Darwin. Not much of a mathematician if you can't."

He succeeded in giving at least his junior students (including myself) the impression that his own mathematics was very lowly indeed. He would work out a simple, Higher-Certificate-level, problem on the dynamics of impact on the blackboard, at great length and with an appearance of immense pride at having actually done it all by himself, and turn to us at the end, saying: " Now that 's the sort of thing every physicist ought to be able to do for himself." This gave his students the impression that their professor could just about do something they had themselves been able to do quite easily for years and years. It is perhaps worth mentioning that at the university at Christchurch, New Zealand, in 1893 he took his finals in both mathematics and physics, and in each of these subjects he got the only first awarded in the university in that year.

Every one can be encouraged by thinking of those two great men, Faraday and Rutherford. They were akin in enjoying almost every moment of life, in having little liking for very abstruse mathematics or for experiments needing complicated apparatus, and of having a visual imagination of such superb clearness that each could see, at every point in his scientific journey, where to go from here.

It was said at Cambridge that one could gauge the progress of a Trinity feast by standing on the far side of Great Court (about 100 yards across) and listening to Rutherford laughing inside the hall. Whenever, to any student who feels himself to be a bit of a fool, a formerly difficult bit of atomic physics suddenly becomes obvious, he should be able to catch an echo of Rutherford's laugh.

The first artificial transmutation of an element was done by Rutherford at Cambridge in 1919. By bombarding nitrogen with α-particles (that is, with stripped helium nuclei emitted by radioactive substances at speeds of the order of one-tenth of the speed of light) he obtained a proton, or hydrogen nucleus, and an oxygen nucleus with the unusual atomic weight of 17 instead of the usual 16.

What happened may be represented by the nuclear reaction:

$$ {}_2\mathrm{He}^4 + {}_7\mathrm{N}^{14} \rightarrow {}_8\mathrm{O}^{17} + {}_1\mathrm{H}^1. $$

The subscripts give the atomic number, which is not only the order of the element in the periodic table, but also the positive charge on the nucleus in electronic units.

The superscripts give the atomic weight in mass units.[1] Thus $_7N^{14}$ represents a nitrogen nucleus, seventh in the periodic table and having an atomic weight of 14 mass units. $_8O^{17}$ represents the nucleus of a rather rare isotope of oxygen, eighth in the periodic table and weighing 17 mass units.

In the equations for nuclear reactions the conservation of charge and mass requires that both the superscripts and the subscripts add up to the same on both sides.

In the example given we have

for the superscripts $4 + 14 = 17 + 1$,
and for the subscripts $2 + 7 = 8 + 1$.

If we assume, as is generally believed, that each nucleus is made up of protons, $_1H^1$, and neutrons, $_0n^1$, then we can see at once that $_7N^{14}$ is made up of 7 protons and 7 neutrons, and $_8O^{17}$ of 8 protons and 9 neutrons.

This small-scale experiment of Rutherford's may be regarded as the first major step of the human race either downwards to destruction or upwards to the possession of all the power (in the strict physicist's sense of rate of production of available energy) that any one could need for any reasonable purpose.

Rutherford's nuclear theory of the atom was one of the major steps in the direction of the atomic bomb. The other was taken some years earlier, in 1905, when Einstein published a paper on what we know as the Special Theory of Relativity. One of the consequences of this theory was the relationship $E=mc^2$, a statement that a mass of m grammes is actually equivalent to E ergs of energy, c being the velocity of light. Thus 1 gramme of mass was shown to be somehow convertible into the immense amount of energy represented by 9×10^{20} ergs. The remote theorizing of Einstein, Planck, and Rutherford before 1914 led in 1945 to the horror of Hiroshima and Nagasaki.

It may seem odd to say first that Rutherford's transmutation experiment led to the atomic bomb and then that Einstein's mass-energy equation led to the atomic bomb, but this is the fact. Both the experiment and the apparently unrelated idea were needed to make nuclear energy available on a large scale. We can make the meaning of the equation precise by working out an example. In the equation $E = mc^2$, E is the energy in ergs liberated when m grammes of matter are

[1] A full explanation of unfamiliar terms must wait for Chapters IV and VI. See Index.

dissipated as energy; c is the velocity of light, 3×10^{10} cm./sec. Thus, if 1 gramme of matter is dissipated, the energy released is given by

$$\begin{aligned} E &= 1 \times (3 \times 10^{10})^2 \text{ ergs} \\ &= 9 \times 10^{20} \text{ ergs} \\ &\doteq 21{,}400{,}000{,}000{,}000 \text{ calories.} \end{aligned}$$

This is just about the amount released by the atomic bomb over Hiroshima, or by detonating 20,000 tons of T.N.T., or by burning 2,500 tons of coal, which releases about 8,000 calories per gramme.

From 1919 to 1932 the major advances were theoretical, in the quantum mechanics of Heisenberg, Born, and Jordan, and the wave-mechanics of de Broglie, Schroedinger, and Dirac. Through this theoretical work the puzzle of the relation between the wave and quantum theories of light began to be solved.

De Broglie was the first to suggest the extremely odd relationship between matter and radiation which has since formed the basis of a successful theory of the spectra of atoms with more than one electron (Bohr's 1913 theory having worked with complete success only with single-electron atoms such as hydrogen, singly-ionized helium, doubly-ionized lithium, and so on). De Broglie advanced the theory in 1925, and in the same year Elsasser suggested that electrons might be diffracted by metals as X-rays are by crystals. This was shown to be a fact by Davisson and Germer in 1927 with electrons reflected from the surface of a metallic crystal, and G. P. Thomson, son of J. J. Thomson, gave a still more striking demonstration in 1928 by passing electrons through thin metallic plates. He got a central spot surrounded by diffraction haloes, just like those which a beam of homogeneous X-rays might have given, but corresponding to a shorter wave-length. A magnetic field was found to deflect the ring system as a whole, showing conclusively that the diffraction pattern was due to electrically charged particles and not to electromagnetic waves. In the same year Kikuchi in Japan got similar results with a thin mica crystal.

These experiments forced physicists to believe, however reluctantly, in the wave-like properties of matter, just as the photoelectric effect had forced them, over twenty years earlier, to believe in the matter-like properties of electro-magnetic waves.

After 1920 the wide development of broadcasting led to a great diffusion both of the knowledge and of the technique of very high-frequency oscillatory currents, whose behaviour differs from that of ordinary varying currents. Broadcasting also led to the investigation of the Heaviside layers (so called after the great mathematician Oliver Heaviside), the conducting layers in the atmosphere high above the earth. Trains of wireless waves can be partially reflected from these layers —it is this partial reflection which prevents all the energy radiated from a broadcasting station from escaping into space and makes direct long-distance communication possible—and the timing of the return signals gives an estimate of the height of the reflecting surface. As part of the signals penetrates the layer—the more nearly normal the incidence the larger the proportion which penetrates, as with light—signals can be sent to the moon. Echoes are found to get back in about $2\frac{1}{2}$ seconds, and the distance covered by electric waves in $1\frac{1}{4}$ seconds ($1\frac{1}{4} \times 186{,}000$ miles $= 232{,}500$ miles) is about the lunar distance. In 1925 Millikan did the first quantitative measurements of the mysterious cosmic rays which reach the earth from outer space.

In 1932 came an immense burst of discovery. Cockcroft and Walton for the first time disrupted an atomic nucleus with artificially accelerated particles instead of using α-particles, and E. O. Lawrence in America constructed the cyclotron to do the same job. The point about artificial as against natural disruption is that you can do a lot more of it. With natural α-particles you could never get enough matter transmuted to do more than show you had done the trick; but with artificial projectiles you could get enough of the material to enable you to see it.

The result of the 1932 illumination was the discovery of three new particles—the neutron, the positron, and the neutrino; or, to be more exact, the discovery of the first two and the inference of the third. In 1936 the meson was first observed, and in 1939, just before the second great war of the twentieth century, came the discovery of the fission of uranium and the availability of really large amounts of nuclear energy; and hence of the atomic bomb.

These discoveries, from 1932 onward, are discussed in Chapter VI.

CHAPTER II

ELECTRONS AND POSITIVE RAYS

Crossed Fields—Parallel Fields—Finding $\frac{e}{m}$ for Electrons. Thomson, Dunnington—Finding e. C. T. R. Wilson, Townsend, Thomson, H. A. Wilson, Millikan, Perrin—Positive Rays. Goldstein, Thomson, Aston, Dempster, Bainbridge—The Discharge-tube and Macleod Gauge—Ions in Gases at Atmospheric Pressure—Mobility of Ions.

THIS has been an enormous subject. It filled the arena of research in physics with blood, toil, tears, sweat, entertainment, and friendship from 1894 till 1919. Students to whom it is quite new are advised to start with the section The Discharge-tube and Macleod Gauge on p. 527, in order to get a general idea of the problems being tackled, and the way they appeared to the early investigators. The discharge-tube experiments should be seen. The objection to putting the discharge-tube section at the beginning of the chapter is that it leads directly to very complex phenomena which are too difficult for the beginning of the chapter. Students who start with an elementary study of the discharge-tube should return to it in its proper place in the chapter. Among the greatest names in the investigations of e and $\frac{e}{m}$ for electrons and positive rays are J. J. Thomson, Townsend, C. T. R. Wilson, H. A. Wilson, Millikan, and Aston.

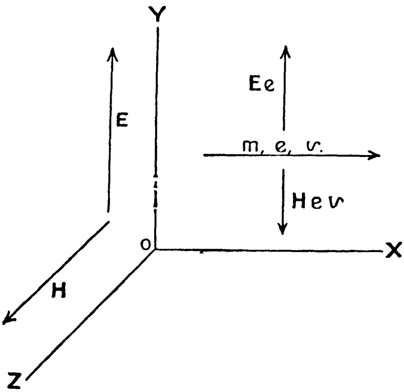


FIG. 242.

Instead of describing all the methods and experiments separately, let us begin with two simple problems: the movement *in vacuo* of a charged particle (charge e, mass m) at right angles to an electric and to a magnetic field when these fields

are crossed (i.e. at right angles to each other) so that their effects on the charged particle act in opposite directions, as shown in Fig. 242; and the corresponding movement when these fields are parallel.

Crossed Fields

Suppose the charged particle is moving outwards parallel to the X-axis, the electric field E is acting parallel to the Y-axis, and the magnetic field H is parallel to the Z-axis. Then the force Ee due to the electric field acts parallel to the Y-axis, and the force Hev due to the magnetic field also acts parallel to the Y-axis but in the opposite direction, as can be seen by Fleming's left-hand rule for the force on a current-element ev at right angles to a magnetic field H.

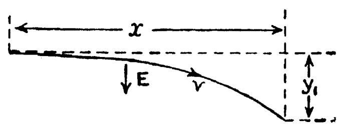


FIG. 243.

A charge e moving with velocity v is equivalent to a current-element ev of length 1 cm. (see p. 346). If the forces due to E and H exactly counterbalance one another, so that the moving charge keeps on in a straight line, Ee = Hev, and hence we have $v = \frac{E}{H}$, a simple and useful result.

Now let us suppose that over a certain range both E and H can be made uniform, and that they can be made to stop abruptly at the end of that range. Then suppose that the length of travel within the uniform field is x, and the deflection at the end of it is y_1 for the electric field alone, and y_2 for the magnetic field alone. For y_1 see Fig. 243. For y_2 see Fig. 244.

Let us consider the electric field alone first. Since y_1 is very small compared with x, we may regard it as a short distance covered from rest in time $t\left(=\frac{x}{v}\right)$ with an acceleration $\frac{Ee}{m}$ due to a force Ee acting on a mass m.

We have then $$y_1 = \frac{1}{2}\frac{Ee}{m}t^2 = \frac{Eex^2}{2mv^2}$$

and so $$v^2 = \frac{Eex^2}{2my_1}.$$

For the magnetic field alone the curve described by the particle will be circular, for the field H acts perpendicularly to the plane of the paper, and the force on the particle is always perpendicular to its direction of motion. Let the radius of the circular path described be r.

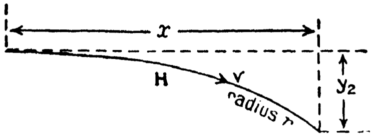


FIG. 244.

From the geometry of the circle, applying Pythagoras' theorem, we see that

$$r^2 = (r - y_2)^2 + x^2$$

and hence $$r = \frac{x^2 + y_2^2}{2y_2}, \qquad \text{which gives us } r.$$

Since the acceleration toward the centre is equal to the force

$$\frac{mv^2}{r} = Hev, \text{ and so } v = \frac{Her}{m}$$

and

$$v^2 = \frac{H^2e^2r^2}{m^2}.$$

Equating the two values of v^2, we get

$$\frac{H^2e^2r^2}{m^2} = \frac{Eex^2}{2my_1}$$

and hence

$$\frac{e}{m} = \frac{Ex^2}{2H^2y_1r^2}.$$

All the quantities on the right-hand side of this equation can be observed or calculated from observation. We could also have found $\frac{e}{m}$ from either the magnetic or the electric deflection, alone, by putting instead of v the balancing values of $\frac{E}{H}$, which kept the particle going straight without deflection.

For the electric deflection we have

$$\frac{e}{m} = \frac{2y_1v^2}{Ex^2} = \frac{2y_1}{Ex^2}\cdot\frac{E^2}{H^2} = \frac{2y_1E}{H^2x^2}.$$

For the magnetic deflection

$$\frac{e}{m} = \frac{v}{\mathrm{H}r} = \frac{\mathrm{E}}{\mathrm{H}^2 r}.$$

These methods of working out the results are, however, ideal, not practical. Neither the electric nor the magnetic field is ever so manageable. In practice, $v = \frac{\mathrm{E}}{\mathrm{H}}$ can be taken as merely giving a rough indication of the speed, and the fields E and H have to be laboriously explored from point to point, and their cumulative effect found by approximate graphical methods. Since we may take $\frac{1}{r} = \frac{d^2y}{dx^2}$ as true enough for the small curvatures involved, we have for the magnetic field

$$\frac{d^2y}{dx^2} = \frac{\mathrm{H}e}{mv},$$

so that

$$y_2 = \frac{e}{mv}\int_0^x\left[\int_0^x \mathrm{H}dx\right]dx,$$

and for the electric field, by differentiation of the equation

$$y_1 = \frac{\mathrm{E}ex^2}{2mv^2}$$

(which is true for each point but cannot be used as it is all the way along because of variations in E),

$$\frac{d^2y}{dx^2} = \frac{\mathrm{E}e}{mv^2}.$$

So that if we want an equation which *is* true for the length x as a whole we get it by integrating, and it is

$$y_1 = \frac{e}{mv^2}\int_0^x\left[\int_0^x \mathrm{E}dx\right]dx.$$

These equations for y_1 and y_2 enable us to find $\frac{e}{m}$ and v when the distribution of E and H is known.

Parallel Fields

Let us now consider the effect when E and H are not crossed but act in the same line (perpendicular to the direction

of motion of the particle, of course). In this case the forces acting on the particle are in two directions at right angles to each other as well as to the direction of motion of the particle.

The situation now is shown in Fig. 245. E and H both act parallel to the Z-axis, and so does the electric force Ee on the particle.

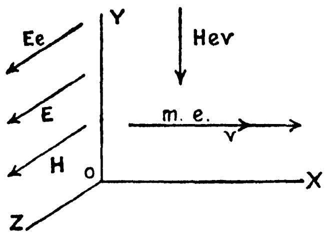


FIG. 245

The particle itself is travelling parallel to the X-axis, and hence, by Fleming's left-hand rule, the force Hev due to the magnetic field acts parallel to the Y-axis in a negative direction for a positive value of the charge e. If, as before, the particle covers a distance x at velocity v in time t, then $x = vt$. During this time it moves a distance y parallel to the Y-axis owing to the field H, and a distance z parallel to the Z-axis owing to the field E.

We have then, as before,

$$y = \frac{1}{2}\cdot\frac{\mathrm{H}ev}{m}\cdot\frac{x^2}{v^2} = \frac{\mathrm{H}ex^2}{2mv}.$$

$$z = \frac{1}{2}\cdot\frac{\mathrm{E}e}{m}\cdot\frac{x^2}{v^2}$$

Eliminating v from these two equations we get

$$v^2 = \frac{\mathrm{E}ex^2}{2mz} = \frac{\mathrm{H}^2e^2x^4}{4m^2y^2}$$

so that

$$y^2 = \frac{x^2\mathrm{H}^2}{2\mathrm{E}}\cdot\frac{e}{m}z.$$

Now if a screen is erected so that the X-axis is a normal to it, at a point so that the particle hits it at the end of the time t, it follows that the particles impinging on it will hit it along a curve whose equation is that just given. This is of the form $y^2 = \mathrm{K}z$, and is, therefore, a parabola. The particles will be distributed along this parabola according to the values of v they happened to start with, and, if x, E, and H are known, the whole parabola may be used to find a value of $\frac{e}{m}$.

This calculation is the basis of J. J. Thomson's famous Positive Ray Parabolas.

Finding $\frac{e}{m}$ for Electrons. Thomson Dunnington

The first arrangement we have described, that of crossed fields giving directly opposing deflections, was used for J. J. Thomson's method of finding $\frac{e}{m}$ for the cathode particles in a discharge-tube, the particles which in this experiment he identified as what we now call electrons; and variations on part of this arrangement were devised by Kaufmann and also by Wiechert in their independent determinations of $\frac{e}{m}$.

Thomson's method can be easily repeated by students in an advanced laboratory (see Worsnop and Flint); Kaufmann's

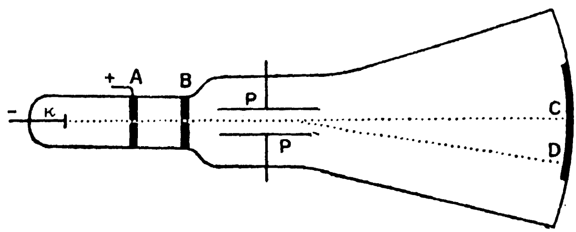


FIG. 246.

gave far the most accurate result at the time; Wiechert's measured the speed of the cathode rays directly and found it about one-tenth of the speed of light, and so proved directly that they were particles, not electromagnetic radiation.

Fig. 246 shows Thomson's method diagrammatically. Electrons leave the cathode K and penetrate the perforated anode A. They then pass through a hole in a screen B whose purpose is to narrow the beam and keep it as near a thin pencil as possible. This pencil then passes between metal plates PP which can be electrified, and at the same time between poles of an electromagnet which may be imagined above and below the paper. If neither field is on they strike C in the centre of a screen of fluorescent material at the end of the tube. If either field is on, they are deflected to some other point

on the screen, represented by D. The length CD gives the value of $\frac{e}{m}$, as we have shown.

The most accurate method (up to 1949) is Dunnington's, which is a modification of one devised by Wiechert in 1899. Electrodes from a cathode C (a hot filament) are accelerated through four slits in succession, S_1, S_2, S_3, S_4, situated on the arc of a circle. The first slit S_1 is also an anode and carries the accelerating potential, which is an alternating one, so that the electrons travel in successive pulses.

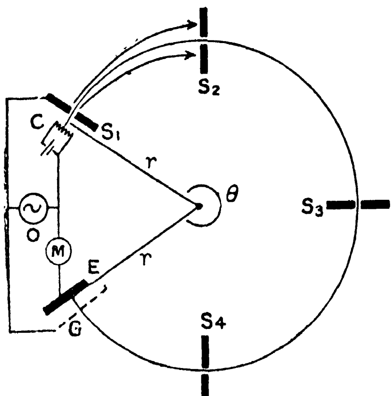


FIG. 247

A magnetic field H, of great uniformity and necessarily very accurately known, acts vertically downwards through the plane of the circular path, to keep the electrons in their circle. Faster and slower electrons than those covering a very narrow range of velocities are caught in the side-pieces of S_2. The survivors reach the collecting plate E, unless stopped by an adverse potential at the grid G. An oscillator O applies the alternating E.M.F. to the cathode C and (through a current-measuring instrument M) the collecting plate E on one side, and the anode S_1, and the grid G on the other side.

The electrons will only get to E if the time they take to get from their effective starting point round to E (through the whole angle θ) is *not* any multiple C of the period T of the oscillator. If it is, they will be stopped by G. Thus the velocity v is given by $v = \frac{r\theta}{nT}$, where r is the radius of the circle described by the electrons.

But we already have proved that

$$Hev = \frac{mv^2}{r}$$

for an electron describing a circle of radius r with velocity v at right angles to a magnetic field H. Thus

$$\frac{e}{m} = \frac{v}{Hr}$$
$$= \frac{r\theta}{nTHr}$$
$$= \frac{\theta}{nHT}.$$

This method gives very good accuracy indeed, having a probable error of less than 1 in 4000. Notice that r does not appear in the final equation.

Dunnington got $\frac{e}{m} = 1{\cdot}7597 \pm 0{\cdot}0004 \times 10^7$ e.m.u./gm.

Note also the immense importance of the $\frac{e}{m}$ experiment. The experimental value found for $\frac{e}{m}$ for cathode rays had nothing to do with either the nature of the filament or the nature of the filling gas. It was the same for thermions, the electrons emitted from a hot metallic surface, see p. 602, and for photo-electrons. It clearly belonged to some universal kind of particle with an invariable charge. If that charge were the same as the charge on the ion in electrolysis, namely the Faraday divided by Avogadro's number, then the mass of the particle could be deduced. It was $\frac{9649}{1{\cdot}76 \times 10^7 \times 6{\cdot}023 \times 10^{23}}$ or $9{\cdot}1 \times 10^{-28}$ gm.; at the time of the early experiments an amazingly small mass, little more than $\frac{1}{2000}$ of the mass of the hydrogen atom, then believed the smallest of all particles, and indivisible.

The cathode-ray oscillograph, already briefly mentioned on p. 445, may be said to have developed from the work on $\frac{e}{m}$ for electrons.

It is shown diagrammatically in Fig. 248. Though it does not use a magnetic field, it does use two electric fields, F_1 and F_2, at right angles to one another. Electrons are accelerated from the cathode C through the perforated anode

A (maintained at something up to 6000 volts positive with respect to the cathode), and they fall on a fluorescent screen SS. When leaving the cathode the electrons are concentrated by passing through a metal cylinder (not shown) called a Wehnelt cylinder. The system of cathode (an oxide-coated filament), Wehnelt cylinder, and perforated anode is called for obvious reasons an electron gun. Usually the first field F_1 has a definite frequency of non-sinusoidal alternations. If it is applied alone the spot of light on S (and very good concentration of the electron-beam is needed to make it a small enough spot) moves across the screen at uniform speed

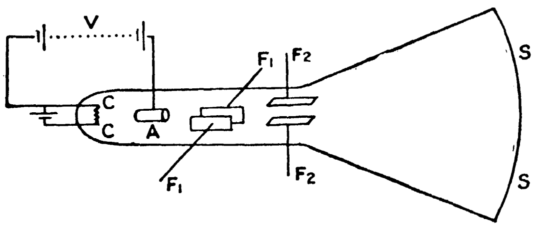


FIG. 248

and then flashes as quickly as possible back to its starting point. The alternating field whose wave-form is to be investigated is then applied through the F_2 plates, and can be analysed through the perturbations it causes. Magnetic fields can also be applied outside the tube, and so investigated.

The point of the cathode-ray oscillograph is that it has practically no inertia because its moving parts are electrons; and so it can be used to measure time intervals as short as 10^{-8} seconds, and study elements lasting 10^{-8} seconds or longer in transient phenomena whose total duration might be of the order of 10^{-6} seconds.

Finding e. C. T. R. Wilson, Townsend, Thomson, H. A. Wilson, Millikan, Perrin

All the early measurements of e, except Townsend's, depend on a principle first developed by C. T. R. Wilson and published in 1897, that year of marvels, in his account of what has come to be known as the Wilson cloud chamber. This was an arrangement for producing a sudden adiabatic expansion in an

unsaturated mass of air, and thus making it suddenly supersaturated. Condensation in supersaturated air happens more readily on charged particles than on uncharged, and more readily on negative ions than on positive. For example, air just saturated will condense on negative ions for a 1·26-fold expansion, and for positive ions for a 1·30-fold expansion.

When Wilson's method first appeared it enabled Thomson to improve on a method already used by Townsend. Townsend had produced an electrified cloud by bubbling a known amount of air through water and charging it with a known amount of electricity. He found the radius of the droplets, and hence the mass of each, by timing the rate v at which the cloud settled down and applying a formula originally evolved by Sir George Stokes. According to this formula, the resistance to a particle of radius a falling through a gas of viscosity η is $\frac{4}{3}\pi a^3 \rho g$, where ρ is the density of the drop. Thus

$$\text{viscous resistance} = \text{weight of drop}$$

$$6\pi a \eta v = \frac{4}{3}\pi a^3 \rho g$$

and we get
$$v = \frac{2}{9}\frac{a^2 \rho g}{\eta}$$

Since the mass of the drop is given by $m = \frac{4}{3}\pi a^3 \rho$, he could find the number of drops in the cloud by dividing m into the mass of the cloud. Knowing the total charge, he could find the average charge per drop. This method did not allow for the fact that many drops carried multiple charges or that the drops varied in size.

Thomson's method was effectively the same as Townsend's, but it used Wilson's cloud chamber, and calculated the weight of the cloud from a theoretical deduction of the cooling produced by expansion and the difference of temperature between the room and the expanded cloud He therefore had Townsend's uncertainties and this new one as well; but his method was capable of far more improvements than was Townsend's, and historically it is therefore the more important of the two. In 1903 H. A. Wilson devised a way of eliminating the worst error, that of the multiply-charged drop. He did this by producing an idea of extreme brilliance. He applied an electric field acting vertically upwards, and succeeded in

holding individual layers of drops stationary in the field of the microscope. If a field E holds the drop stationary, then

$$Ee = mg$$

m was found as before by applying Stokes's Law to the rate of fall of the drops. Wilson found that under the action of the electric field the cloud tended to divide into stratified layers, one presumably with singly-charged droplets and others with multiply-charged droplets.

The errors in H. A. Wilson's work were mainly due to the same inaccuracies in applying Stokes's Law, and to the tendency

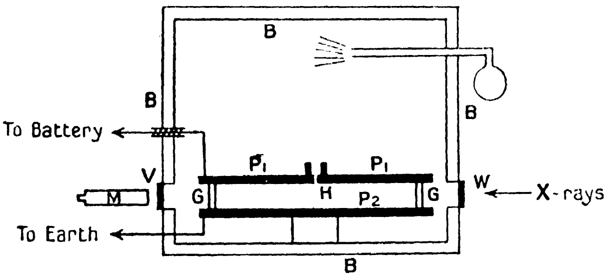


FIG. 249

of the expanded gas to warm up to room temperature, reduce the saturation, and allow individual drops to be evaporated, so that m was liable to alter during the course of the experiment. H. A Wilson's method did, however, show conclusively that some drops had double, triple, or even higher charges.

The inaccuracies were avoided successfully by Millikan, who avoided the evaporation difficulty by using oil-drops.

His apparatus consisted of two optically flat brass plates, P_1 and P_2, held exactly parallel by glass supports GG. P_2 was earthed, P_1 connected to a battery or source of potential giving a voltage of the order of 6000 volts, positive or negative, to earth. The plates were enclosed in an earthed metal box BB. Suspended drops could be watched by a microscope through the window V. They could be illuminated by light from a third window (not shown) in a direction at right angles to the line VW. Drops were usually ionized by friction, but X-rays for ionization could be applied through W if

needed. Drops of oil or mercury could be sprayed through the sprayer S, and would sink, one or two at a time, through the small hole H in the top plate. They might either get charged by friction in the sprayer, or by ions caused by the *X*-rays between the plates. The whole apparatus was kept in a thermostatic enclosure to avoid errors from convection currents. Evaporation was avoided partly by this thermostatic control and partly by using oil or mercury, which do not evaporate appreciably in such conditions. The chief difficulty was in the application of Stokes's Law to find the mass of the suspended drop, and Millikan took immense trouble to overcome this.

He allowed for the density of air as well as the density of the drop by taking $\frac{4}{3}\pi r^3(\rho - \sigma)$ as the effective mass of the drop, ρ being the density of the drop and σ the density of air. He investigated the accepted value of the viscosity of air, and found it about 2% wrong. Instead of taking Stokes's accepted formula $6\pi a\eta v$ as being the viscous resistance to a sphere of radius a, in a gas of viscosity η, he made a correction to allow for the fact that the length of the mean free path of air molecules is not really negligible compared with the size of the drop. (Millikan expressed this by saying " the radius of the drop has begun to be comparable with the mean size of the holes " in the air.) $6\pi a\eta v$ became $6\pi a\eta v \left(1 + \frac{A\lambda}{a}\right)^{-1}$, where λ is the mean free path and A a constant, he determined by working with many different sizes of drop.

In the course of his long series of experiments—he was hard at work at cosmic ray experiments during 1948 at the age of eighty and has returned to the measurement of e several times since his first paper in 1910—Millikan settled a number of vexed questions. He found that charges on positive and negative ions are invariably multiples, within the limits of experimental error, of the same unit, whether the charge was due to collisions, or *X*-rays, or friction in the sprayer. Usually the charge carried on a drop from the sprayer was between 3 and 20 times the unit. On a gaseous ion the charge was usually single, though in helium ionized by α-particles about 1 ion in 6 carried a double charge. Millikan's 1923 value for e was $4{\cdot}770 \times 10^{-10}$ e.s.u.

One more method of finding e was devised by Perrin. He deduced Avogadro's number from observations on the Brownian movement of particles in a suspension. He found their density distribution at different levels. This interesting method seems to have involved too many experimental errors, for it gave much too high a value of N, and a better way of finding N from the Brownian movement by following the rate of diffusion of individual particles has been devised by Einstein and was carried out by Fletcher with a result coming very close to Millikan's. The value of e is found by dividing N into the Faraday. Perhaps the greatest achievement of these e measurements was the conclusive demonstration of the discontinuity of charges, the atomic nature of electricity.

Positive Rays. Goldstein, Thomson, Aston, Dempster, Bainbridge

Positive rays were discovered by Goldstein in 1886 by making a hole in the cathode of the discharge-tube. Wien demonstrated that they were indeed positively charged particles by observing the effect on their motion of an electric or a magnetic field. Goldstein called the hole in the cathode a canal, and hence these rays were called *Canalstrahlen* by Goldstein and Goldstein rays by others.

Thomson set out to investigate these rays, and began to publish his work in 1910. The method he had used for cathode rays was relatively inaccurate, for the immensely greater mass of the particles made the deflections much smaller. He therefore devised a quite different way of measuring $\frac{e}{m}$: the method known as "positive ray parabolas."

The apparatus is shown diagrammatically.

B is an ordinary discharge-tube bulb of about 1500 c.c. capacity, with anode A, and a connection (shown at the top) to an exhausting pump. Another connection G allows the gas under examination to be admitted through a very fine tube. The cathode K is about 7 cm. long, usually of aluminium, penetrated by a very fine copper or brass tube TT, with an internal diameter from 0·1 to 0·01 mm.

WW is a water-jacket for cooling. I is an iron tube (longer than shown) for shielding the whole length of the cathode from external magnetic effects which might make the

rays hit the side of the tube. N and S are the poles of a big electromagnet. EE, EE are soft iron plates insulated from the electromagnet by mica sheets MM, MM.

These plates carry the electric field. As they are soft iron, they also carry the magnetic field. This extremely ingenious arrangement is possible because in this apparatus, unlike the apparatus for finding $\frac{e}{m}$ for cathode rays, the electric and magnetic fields act together in the same direction. Their deflections are therefore at right angles to one another, as we

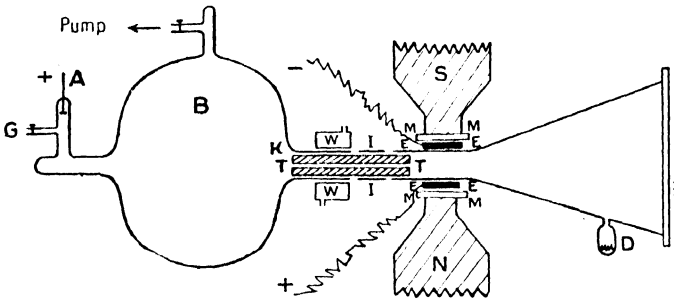


FIG. 250

have shown at the beginning of this chapter; and, as we there showed, the rays falling on the photographic plate P form a series of parabolas, each parabola characteristic of a particular value of $\frac{e}{m}$. D is a drying-tube, probably containing P_2O_5.

As the particles travel to P in straight lines after leaving the field district, those which appear on P are simply enlargements of what would have been found at the end of the fields.

It was by these parabolas that neon was first shown to be a mixture of two isotopes of atomic weight 20 and 22 units, in such proportion as to give a chemical, or average, atomic weight of 20·2. This should have been at the time an earth-shaking discovery, for it suggested that Prout's 1815 guess had been correct, and that all nuclei were built up of complete units, each of a mass very near that of the hydrogen nucleus. Thomson, however, was so much surprised by the result that

he put it forward tentatively, and the definite announcement was made later by Aston.

The next big advance was Aston's mass-spectrograph.

The scheme of this is simple. A stream of positive rays from a discharge-tube comes through narrow slits SS and then passes between charged plates PP. These divert the slow-moving particles more than the fast ones. The streams now pass through a magnetic field M, perpendicular to the electric field. This diverts the slow-moving particles again

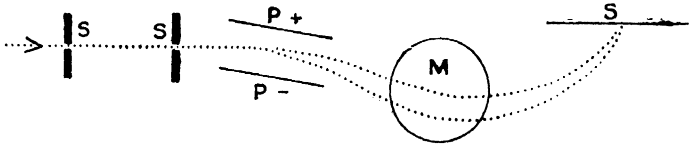


FIG. 251.

more than the fast ones, and thus, with suitable design of apparatus, can be made to bring all particles of a particular $\dfrac{e}{m}$ to a focus on the screen.

Aston's method is used not to give exact values of $\dfrac{\text{charge}}{\text{mass}}$ from first principles, but to compare values with a standard. At this its success has been amazing. Ions differing by only 1 % in $\dfrac{\text{charge}}{\text{mass}}$ can be made to come to a focus 6 mm. apart, and the focus is sharp enough to separate marks 0·005 mm. apart on the film. Thus differences of $\dfrac{0{\cdot}005}{6} \times \dfrac{1}{100}$, or $\dfrac{1}{120{,}000}$, of the atomic mass can be detected.

Dempster's method, though less accurate than Aston's, has the advantage of working easily with metals which (except for mercury) are difficult to introduce as gaseous ions into Aston's mass-spectrograph.

A hot wire FF, heated by the cell V_1, emits electrons which are accelerated by the battery V_2 on to a salt of the metal to be investigated, on the block M. This salt is thus made to emit positive rays, some of which pass through the slit S_1 with low velocity, where they are accelerated by V_3, about 1000

volts, toward the slit S_2. Some, in a thin pencil of ions of almost uniform velocity, pass through this, and are bent through a semicircle by a uniform field H set up by a large electromagnet (not shown). This field H may be considered as acting vertically upwards through the paper. The left-hand Rule shows that positively charged particles travelling in the direction $S_1 S_2 S_3 S_4$ will be bent into a circular path in the direction shown.

The thin pencil passing through S_2 will be kept thin by

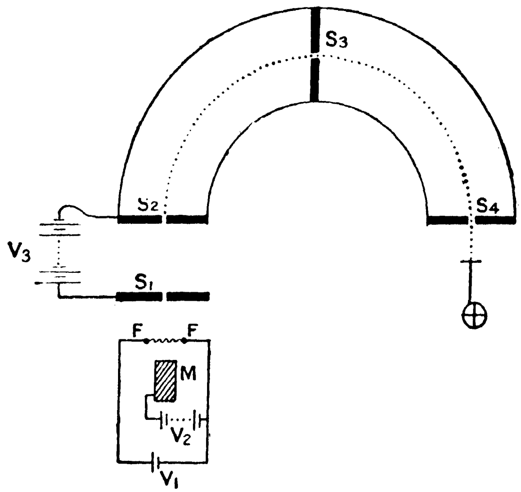


FIG. 252.

passing through S_3 and S_4. The survivors are caught by an electrometer when they emerge from S_4. The current caught by the electrometer is plotted against variations in V_3, and a series of peaks is obtained, each corresponding to a particular value of $\frac{e}{m}$ for the particles entering S_1. As with the mass-spectrograph elements whose mass is not accurately known are often tested against known ones, but absolute determinations can be made.

If r is the radius of the circle $S_2 S_3 S_4$, we have $Hev = \frac{mv^2}{r}$

for a particle of charge-mass ratio $\frac{e}{m}$ entering S_2 with velocity v, since Hev is the force on the current-element ev and $\frac{mv^2}{r}$ is the acceleration toward the centre, caused by this force. Assuming that the whole of v is due to the fall through the potential difference V_3, we also have

$$V_3 e = \frac{1}{2} Mv^2.$$

Equating values of v from these two equations, we get

$$\frac{2V_3 e}{m} = \frac{H^2 e^2 r^2}{m^2}$$

so that

$$\frac{e}{m} = \frac{2V_3}{H^2 r^2}.$$

This method gave useful peaks at 6·0 and 7·0 atomic mass units for the metal lithium, of chemical atomic weight 6·94.

A method introduced by Bainbridge is as follows:

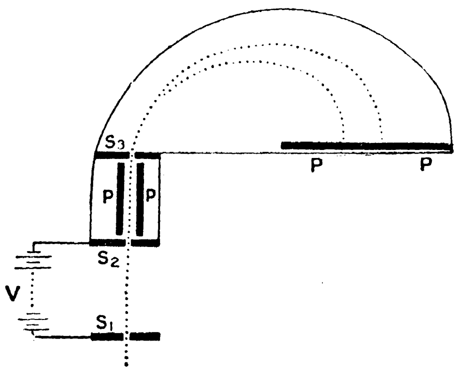


FIG. 253.

Positive ions from a source not shown enter the slit S_1 and are accelerated by a fixed P.D. of V volts (several thousand). They then pass along a narrow channel between slits S_2 and S_3 with electrified plates PP and magnetized plates (not shown), producing electric and magnetic fields E and H_1 at right angles to one another, exactly as in Thomson's apparatus for finding

$\frac{e}{m}$ for cathode rays. Here the apparatus is just a filter. Only particles of a special velocity get through, those for which

$$H_1 ev = Ee$$

and hence

$$v = \frac{E}{H_1}$$

as shown at the beginning of this chapter.

When they leave S_3 the particles enter a uniform magnetic field H_2 (apparatus not shown, but H_2 acting vertically upwards through the paper). Each kind of particle describes a circle of characteristic radius and finishes on the plate P.

Suppose that a particle of mass m_1 and charge e describes a circle of radius r_1, then we have

$$H_2 ev = \frac{m_1 v^2}{r_1}.$$

Dividing through by v, putting $v = \frac{E}{H_1}$, and rearranging, we get

$$\frac{e}{m} = \frac{E}{H_1 H_2 r_1}.$$

This method seems to combine all the advantages. It has more than the mathematical certainty of Thomson's $\frac{e}{m}$ method, the accuracy of Aston's mass-spectrograph (or rather even more accuracy), and the ability to use metallic sources as Dempster's method does, or gaseous sources like Aston's. Modern types of mass-spectrograph have been developed from Bainbridge's method for two distinct purposes. One type gives close determinations for the mass of a particular isotope; the other finds the relative abundance of different isotopes of the same element.

The Discharge-tube and Macleod Gauge

The discharge-tube is so important historically, so common in laboratories, so spectacular, and so amusing to play with, that one tends to forget that one cannot do very much with it except look at it, and that its apparently simple and definite phenomena are not quite fully explained even yet, so that it succeeds in being at once too elementary and too advanced

for the beginner in atomic physics. It does, however, illustrate principles and give one a very strong sense of the reality of electrons and ions.

In the apparatus shown, ABK is the discharge-tube, A being the anode and K the cathode. I is an induction coil. T is a well-greased gas-tight tap. LMNP is a Macleod gauge, which gives the easiest way of measuring low gas pressures. VR is a Töpler mercury exhausting pump. This is troublesome to set up and is an exhausting pump indeed. It may be replaced by an ordinary motor suction pump; but in the event of poverty, breakdown, or fuel cuts, it is a useful thing to have, and I prefer to have one permanently connected to the laboratory demonstration discharge-tube as well as a motor pump.

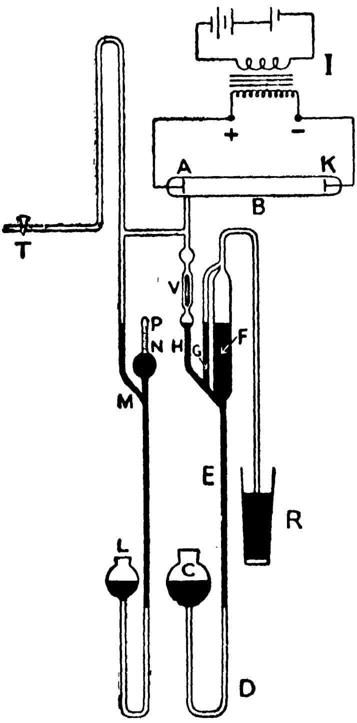


FIG. 254.

The positive output terminal of the induction coil I is connected to A, the metal anode in a glass discharge-tube B. The negative terminal of I is connected to K, the metal cathode in the tube B.

AK is usually about 20 cm., and the tube is about 3 cm. in diameter.

The apparatus below and on the right is a mercury pump for exhausting the tube B.

The open funnel, or cup, C is attached to a flexible rubber tube D, which is attached to a glass tube E, which opens out into three tubes F, G, H. F is widened, and G is narrow. H goes through a one-way valve V to the original tube B. F and G rejoin and go round to a reservoir R.

If C is lowered, the mercury level in F, G, and H falls. When this level is just below the join of F and G, the gas in B expands so as to fill the wide tube F. If C is now raised, the mercury level rises in F, G, and H.

By rising in F, it pushes the gas up and round through the mercury in the reservoir R, from which it escapes into the air.

The level in H rises until it is stopped by the valve V, consisting of a little loose piece of glass rod, ground to fit accurately into a narrow neck just above its normal resting-place.

Each time C is lowered and raised, more gas is taken from B and pushed out of the apparatus.

The measurement of pressure by the Macleod gauge is done as follows. If the cup L is lowered till the mercury level falls below the junction M, then the pressure in the bulb N is the same as that in the apparatus. P is a long narrow graduated closed tube extending above N.

When L is raised again, the gas in N is pushed into P and compressed. If L is raised to a known level, the pressure of the gas in P can be calculated easily from the atmospheric pressure. If the volume of gas in P is known from the graduations, and the volume of N is known originally, then the former pressure of the gas in N can be easily calculated.

The following are the main stages the discharge goes through as the pressure is reduced. The colours described are for air.

Stage 1. Pressure values irregular.

A long thin spark passes between the electrodes A and K. Air always has a few ions in it. These are speeded up by the electric field due to the induction coil.

In the ordinary way, they have so many collisions, losing their momentum each time, that they never get enough speed to ionize the molecules they hit. As the pressure gets less, the molecules get further apart, and the average distance they can go without a collision gets bigger. This average distance is called the mean free path. So there comes a point when they knock enough electrons off the molecules they hit to cause continuous ionization, which shows as a spark.

Atoms or molecules give off light when, and only when, some disturbance takes place in their outer lot of electrons. Light can be emitted without ionization, though ionization is bound to produce light.

If a gas is luminous, it always means that its atoms or molecules are getting knocked about, either by collisions or by electric waves.

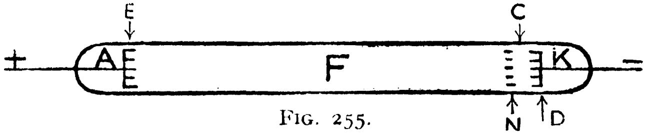


FIG. 255.

Stage 2. Pressure about 3 mm. of mercury. See Fig. 255.

A is the anode.

E is a velvety pink layer of luminous gas near the anode.

F is a long dark space, the " Faraday dark space."

N is a pale violet glow, the " negative column."

C is a short dark space between D and N, the " Crookes dark space."

D is an orange-red velvety glow near the cathode.

K is the cathode.

In the succeeding stages only new phenomena will be given names. The spacing of those already named is shown in Fig. 255.

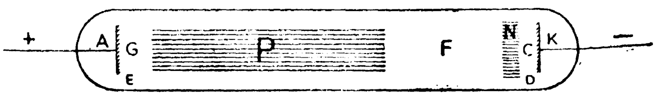


FIG. 256.

Stage 3. Pressure about 1 mm. See Fig. 256.

The negative column N moves out and gets larger.

A long coloured column P, the "positive column," appears going more than half-way down the tube from near the positive end.

A new short dark space G (unnamed) appears between E and P.

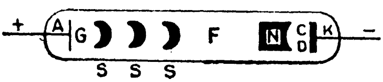


FIG. 257.

Stage 4. Pressure about 0·5 mm. See Fig. 257.

The positive column breaks up into short bits, called STRIATIONS, S S S in the diagram.

Stage 5. Pressure about 0·02 mm. See Fig. 258.

The positive column, after breaking up into striations, gets pushed back toward the anode by the advancing negative column. Finally the positive column disappears.

All this time a green light has been showing on the glass. This is due to electron impact on the glass molecules, which ionizes them, and makes them emit their characteristic light or fluorescence.

As the pressure is reduced further, the negative column is pushed more and more toward the anode, until finally it

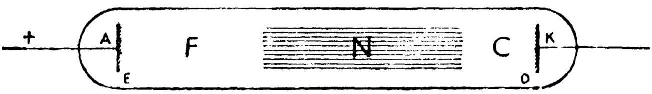


FIG. 258

vanishes, and the Crookes dark space fills the tube. The lower the pressure now gets, the greater is the electrical resistance.

The phenomena depend to some extent on the shape and size of the discharge-tube (or Geissler tube, as it is usually called), the nature of the gas, and the metal of the electrodes. The observations here described were on an air-filled tube about 20 cm. long and 4 cm. in diameter, with aluminium electrodes. One can get the effect with a D.C. steady potential difference of about 1000 volts, but an induction coil is safer and easier to arrange. The positive column is pink and the negative glow blue with air in the tube. The positive column should show the spectrum characteristic of the gas in the tube. The striae should be the more widely separated the wider the tube.

The length of tube occupied by the Crookes dark space is independent of the length of the tube, and so is the length of the negative column, but the positive column has to fit in where it can.

At 0·1 mm. pressure the Crookes dark space is about 1·5 cm. long and the negative column about 12 cm. long, so that at this pressure the tube has to be more than about 14 cm. to show any positive column (or striations) at all.

In the actual experiments here described the writer found the effective resistance of the tube for various pressures

roughly as shown below. (Not resistance in the strict sense, but the quotient $\frac{\text{P.D.}}{\text{current}}$.)

Pressure	*Resistance*
5·0 mm.	500,000 ohms.
2·5 mm.	370,000 ohms.
1·0 mm.	350,000 ohms.
0·5 mm.	500,000 ohms.
0·02 mm.	3,000,000 ohms.

Townsend, *Electricity in Gases*, p. 396, gives the following values for the P.D. required to send a current of 10 milliamps. between aluminium electrodes 11·5 cm. apart in air.

Pressure	*Volts*	*Pressure*	*Volts*
4·0 mm.	650	0·40	530
2·84 mm.	620	0·29	590
1·65 mm.	500	0·24	630
1·04 mm.	470	0·17	740
0·66 mm.	490	0·13	800

These two sets of results agree in finding the conductivity of air greatest at round about 1 mm. pressure.

A rough idea of what is happening is as follows. There are always a few ions in the air, probably produced by cosmic rays or stray radioactivity. An electric field speeds up these ions, and if the pressure is low enough and the molecules in the gas few enough, the mean free path of each ion is long enough for it to get up enough speed between collisions to give it an appreciable chance of causing ionization.

Positive ions go naturally to the cathode, and their impact on it is considered to cause it to emit electrons. (If there is a hole in the cathode, some positive rays stream through it. The direction of their deflection by an electric or a magnetic field shows them to be positive; and a frequently used source for positive-ray work is a discharge-tube cathode with a hole in it). The electrons accelerate away from the cathode toward the anode. It takes them some time to get up enough speed to cause ionization, and the space they traverse before causing it is the Crookes dark space. Ionization happens in the negative column. This contains ions of both signs, but more positive than negative, because on the average the negative ions are more mobile than the positive. The negative column therefore

acts as a positive space charge, and so causes a steep potential gradient (*i.e.* a strong field) in the Crookes dark space, but a very slight potential gradient (*i.e.* a weak field) in the Faraday dark space. Negative ions are slightly speeded up in the Faraday dark space, and the excess of positives in the negative column is balanced by an excess of negatives in the Faraday dark space. There is thus a negative space charge there, which steepens the potential gradient toward the anode end, and speeds up the negative ions enough to cause ionization by collision to begin again, and cause luminosity. This is the beginning of the positive column. The positive column is

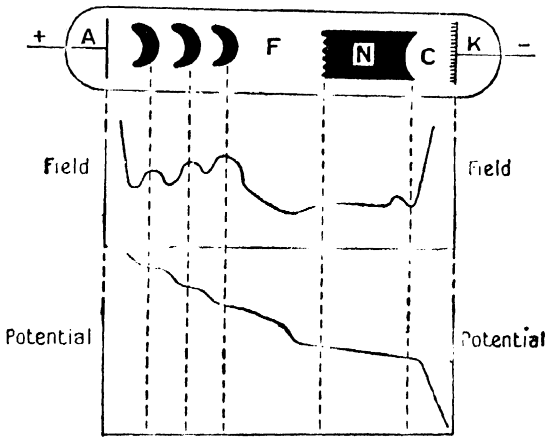


FIG. 259

rather unstable. Accumulations of positive ions may occur as in the negative column. These cause a steepening of the potential gradient on the cathode side and a reduction on the anode side, just as in the negative column. In this way the striae are formed, and each non-luminous part has a steep gradient at the anode end and a slight gradient at the cathode end.

Graham and Wilson measured the field in a discharge-tube by sealing two small electrodes close together in the side of the tube. The anode and cathode were fixed to a frame which could be moved along bodily inside the tube by a magnet, so that the exploring points could be used at any required place

with respect to the anode and cathode. On the assumption that these exploring points do not disturb the potential distribution, the P.D. between them divided by the distance gives the potential gradient, or in other words, the electric field.

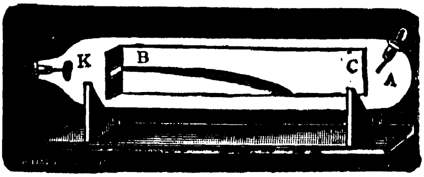


FIG. 260

The electrodes do disturb the potential distribution, and J. J. Thomson got over this difficulty in a most ingenious way. He replaced the exploring electrodes by a stream of electrons going across the tube on to a fluorescent screen on the opposite side. The electric field along the tube displaces this stream of electrons by an amount which gives the intensity and can easily be read off on the fluorescent screen.

The general idea of the distribution of field and potential along a discharge-tube with a pressure of about 0·5 mm. is shown in Fig. 259.

The most instructive result of all is obtained at very low pressure. Before the pressure is very low one can alter the appearance of events inside the tube in a general way by bringing a magnet close to it. The response to a weak magnetic field is enough to demonstrate that much of the discharge is due to the presence of very light moving electrified particles; but at very low pressures it becomes obvious from the direction of the deflection of the stream of particles by a magnetic field that these particles are negatively charged.

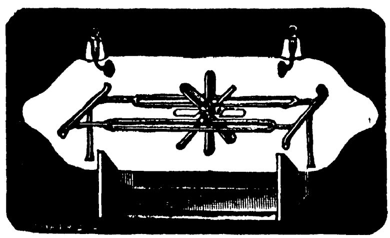


FIG. 261.

When the pressure in a discharge-tube is so low that the Crookes dark space fills it, the electron stream can be made to fall on a fluorescent screen set obliquely across the tube as in Fig. 260. In this the cathode rays from K are made to pass through a narrow slit and fall on the diagonally set fluorescent screen BC.

If the N pole of a bar magnet is brought toward BC from the direction in which we are looking at it, the left-hand rule for the direction of the force on a current in a magnetic field shows that for a current flowing from positive to negative in the direction from C to B, or in other words for a stream of negatively charged particles flowing from B to C, the deflection is downwards as shown.

Other simple demonstrations are given by the apparatus in Fig. 261 and in Fig. 262. In the first the paddle-wheels are made to revolve away from the cathode by the discharge,

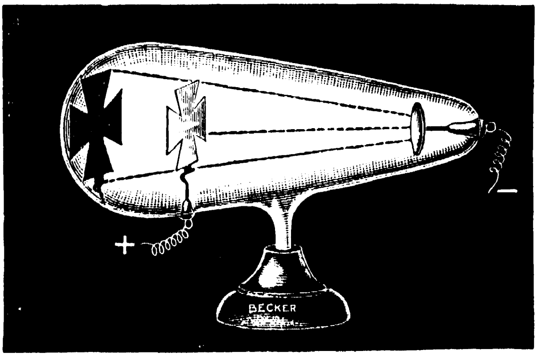


FIG. 262

showing that particles having momentum are coming from the cathode. In the second the Maltese cross, hinged so that it can fall down, casts a shadow at the end of the tube opposite the cathode, and in this shadow no green fluorescence shows. When the Maltese cross falls down, the place where the shadow was shows a brighter green than elsewhere because the glass in the shadow is not fatigued (an interesting but irrelevant effect—the object of the Maltese cross is to show that particles are being emitted from the cathode) and also the shadow can easily be moved about by a magnet. There is still one elementary effect to be demonstrated, the direct deflection of the stream of electrons to one side by an electric field, in a direction which shows that the electrons carry a negative charge.

The electric deflection is much more difficult to demonstrate

than the magnetic deflection. If one puts electrified plates outside the discharge-tube, or introduces a subsidiary electrode into the path of the stream as Perrin did, a very high potential is needed to produce an appreciable effect. Thomson, however, got an effect by passing the stream of particles between parallel plates at low P.D. in a discharge-tube at very low pressure. The difficulty to be overcome is the screening effect of the intensely ionized molecules of the residual gas, and a student is unlikely to succeed in demonstrating the effect for himself.

The purpose of all these simple experiments is to illustrate the following facts unmistakably:

1. In a discharge-tube the cathode is the important electrode. Everything starts from it. The anode is relatively quite unimportant and can be tucked away in a small tube at the side. Its functions suggest those of the male in a matriarchal society; indispensable but inconspicuous.

2. What comes from the cathode is a stream of light electrified particles, not gas molecules, but having some mass and a negative charge.

Ions in Gases at Atmospheric Pressure

Historically, this part of the subject came before the low-pressure work on electrons and positive rays, but it was very puzzling then because so little was known about ions in the less complicated conditions of low pressure. It is still puzzling if one tackles it at the beginning of this part of the subject; and I therefore put it at the end of this chapter.

The flow of electricity through a gas is described by Ohm's Law only in few special sets of conditions; in general it behaves in a much more complicated way. This is because all the the charge has to be carried by ions, charged particles, which except for being charged are almost exactly like the uncharged particles which make up almost the whole population. The number of ions available does not remain constant along the path, but is continually being increased by collision or by stray ionizing radiation, and diminished by recombination. Recombination is combination of positive and negative ions to give neutral atoms or neutral molecules. It is as if a wire carrying a current kept on varying in size, both at individual

places along its length and over its length as a whole, according to the applied potential difference and according to other factors; and as if these changes were often accompanied by luminosity in the wire.

All gases available to us (except perhaps in very special conditions deep down, but not near the bottom, in a lake filled by glacier-fed streams) contain a few ions, caused either by cosmic rays from outer space or by rays from stray radioactive minerals in the earth.

The potential necessary to produce sparking depends on the material and shape of the electrodes, their distance apart, and the nature of the gas between them, as well as on the pressure, the temperature, and the strength of the ionizing agents.

For dry air in normal ionizing conditions at atmospheric pressure, the size of the gap required to produce sparking may be used as a rough guide to the voltage. The following values for breakdown voltages are taken from Holthusen and Braun, 1933.

Gap c.m.	10 *c.m. Terminals* 1000-*volt units*	15 *c.m. Terminals* 1000-*volt units*	25 *c.m. Terminals* 1000-*volt units*	50 *c.m. Terminals* 1000-*volt units*
0·5	17·5			
1·0	32·2			
2·0	60·2	60·7		
3·0	84·1	87·3	88·5	
4·0	105	111	115	
10·0			243	269
20·0				408

One sphere is earthed. The voltage is a little larger if both are insulated.

Ionization is produced in three ways:

1. By direct impact of moving ions, positive or negative, including α- and β-rays.
2. By radiation of suitable (*i.e.* short enough) wave-length such as X-rays, γ-rays, or ultra-violet rays.
3. By heat, as with flames or hot wires.

It can be removed (down to the normal minimum limit) by passing the ionized gas through a wool or cotton plug in a metal tube, or through a long narrow tube, or between charged plates.

Ionization can be studied with an electrometer as shown below.

E is an electrometer, R a high resistance, V_1 and V_2 sources

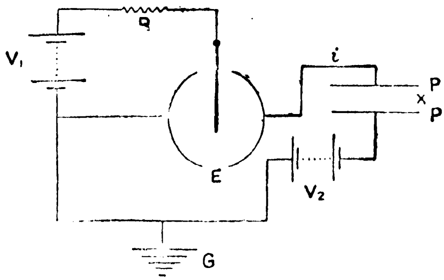


FIG. 263.

of steady potential, PP parallel plates between which is the ionized gas or air being studied, G an earth connection. The current i is measured against the potential V_2, and in general a curve of the type below is obtained:

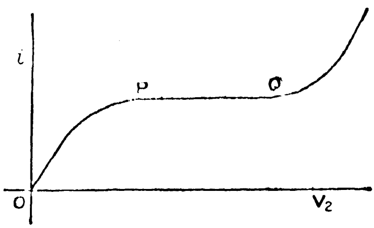


FIG. 264.

The gas is ionized by a source (say a small speck of radium) operating near the point X near the boundary of the plates PP.

What is happening when a current passes between two electrodes in a gas in the presence of an ionizing agent, a permanent potential difference being maintained between the electrodes? (Whether an ionizing agent is specially applied or not, a small ionizing agent is always present, owing to the combined effect of stray local radioactive particles and cosmic

rays. Together they send ionizing particles at the rate of about one per square centimetre per minute.)

Let us suppose the ionizing agent is suddenly brought into action. Ions will be created in pairs, positive and negative, and at first their number will build up, for they will be produced faster than they are lost by recombination, diffusion to the walls of the vessel (small enough to be negligible except in long narrow tubes), and being swept to the electrodes. Then, unless the field is strong enough to speed them up enough between collisions to ionize gas molecules by collision (this will be considered in a later chapter), a steady current will be carried to the electrodes in both directions, and the total number of ions per c.c. of the gas will also remain steady at any given point.

Mobility of Ions

The mobility of an ion is its rate of drift under the influence of unit field. If an iron of mobility k moves with velocity v in a field E, then $v = k\text{E}$. As $k = \frac{v}{\text{E}}$, velocity per unit electric field, it is expressed in cm./sec./volt/cm. (centimetres-per-second per volt-per-centimetre). The mobility is a constant for a given gas under given conditions of temperature and pressure, for an ion of given sign. Positive and negative ions have slightly different mobilities, the mobilities of negative ions being in general greater. The reasons for this would take too long to go into here (but see Stranathan's *Particles of Modern Physics* for a clear and readable account).

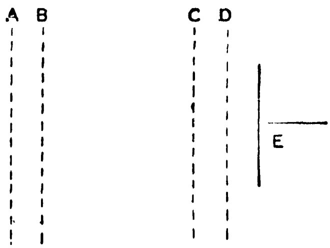


Fig. 265.

Many ways of measuring mobility have been devised, and one of the most powerful is that of Tyndall and Powell.

In principle the arrangement is as in Fig. 265, in which A, B, C, D, are conducting gauzes and E is a collecting electrode.

A small electric field introduces ions to gauze A. An alternating field passes them to B in pulses, one pulse per cycle.

A fixed field between B and C carries them to C. The same alternating field is applied between C and D as was applied between A and B. If the time taken for the ions to go from B to C is a whole number of periods, the pulse will succeed in getting from C to D, and on to the collecting electrode E, which is connected to an electrometer. By varying the frequency of the alternating field one can make the current flowing to E pass through a series of maxima and minima, and finally get a very reliable value for the time the ions take to go from B to C, and hence for the mobility.

Tyndall and his co-workers obtained the following mobilities, measured in cm.-sec. per volt-per-cm. (*i.e.* in units of velocity per unit of electric field) at normal pressure and 18° C.

Ions	*In* He	*In* Ne	*In* N_2	*In* O_2
$Li+$	25·8	11·8	4·2	0·73
$Na+$	24·2	8·7	3·0	0·72
$K+$	22·9	7·2	2·7	0·71

CHAPTER III

THE QUANTUM THEORY. PLANCK, EINSTEIN, BOHR

Planck and the Radiation Problem—Bohr and the Spectrum of Hydrogen—The Beginnings of Wave-mechanics—Critical Potentials in Gases.

Planck and the Radiation Problem

IN the years 1897–9 Lummer and Pringsheim made a series of most careful measurements of the distribution of energy among the wave-lengths of the radiation in a constant temperature enclosure (such as an ordinary oven). Actually they measured the radiation coming out from a very small hole in the oven wall; a hole so small that the radiation issuing from it may be taken as a fair sample of what is inside. It had been proved by Kirchoff in 1859 theoretically, but conclusively, that the radiation in such an enclosure is characteristic only of the temperature and does not depend on the materials in the enclosure. Lummer and Pringsheim were therefore investigating something quite fundamental in the make-up of the material universe.

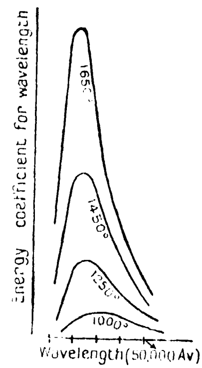


FIG. 266.

Fig. 266 shows the kind of results they got. The four curves shown give the shape of the distribution of energy among the wave-lengths, for absolute temperatures of 1000°, 1250°, 1450°, 1650°. It will be noticed that the wave-length giving maximum energy is about 29,000 Angstrom units (an Angstrom unit, symbol Å, is 10^{-8} cm.) for 1000° Absolute, and about 17,500 Å for 1650° A.; and these figures give the clue to the first fact we should notice about the results. If we multiply the wave-length giving most energy (in centimetres) by the temperature (in Absolute units on the usual scale) we always get about 0·29. Thus $29000 \times 10^{-8} \times 1000 = 0{\cdot}290$ and $17500 \times 1650 \times 10^{-8} = 0{\cdot}289$.

The other two curves shown give similar results.

The second thing to notice is that the energy approaches zero for very long and for very short wave-lengths; and it is on this fact that the whole structure of modern atomic physics may be said to be founded, for all attempts to explain it on the basis of classical physics failed completely.

The development of thought which led to the Quantum Theory began in 1896 when Wien deduced from pure thermodynamical reasoning that the energy per unit volume over the wave-length range λ to $(\lambda + d\lambda)$ in a uniform-temperature enclosure is $E_\lambda d\lambda$ where

$$E_\lambda = \lambda^{-5} f(\lambda T),$$

T being the absolute temperature.

The form of the function $f(\lambda T)$ could not be found from thermodynamics, and in the five years following 1896 a number of attempts were made to find it from other considerations. Of these attempts three are historically important; the last of them, Planck's, being the one which led to the Quantum Theory.

Wien suggested

$$E_\lambda = a\lambda^{-5} e^{-\frac{b}{\lambda T}}\cdot$$

where a and b are constants to be determined. Rayleigh and Jeans proved that the number of degrees of freedom, dn, of a system of waves in an enclosure in the wave-length range from λ to $(\lambda + d\lambda)$ is given by

$$dn = 8\pi\lambda^{-4} d\lambda,$$

and they assumed (without proof) that the average energy associated with each degree of freedom was kT, where k is Boltzmann's constant $\frac{R}{N}$, R being the gas-constant $8{\cdot}315 \times 10^7$ ergs per mole per degree, and N being Avogadro's number $6{\cdot}02 \times 10^{23}$ molecules per gram-molecule or mole. Thus

$$k = \frac{8{\cdot}315 \times 10^7}{6{\cdot}02 \times 10^{23}} = 1{\cdot}381 \times 10^{-16} \text{ ergs per molecule per degree.}$$

Thus $E_\lambda d\lambda = kTdn = 8\pi kT\lambda^{-4}d\lambda$ and hence

$$E_\lambda = 8\pi kT\lambda^{-4}.$$

Lummer and Pringsheim's results were found to agree with Wien's formula for short wave-lengths and low temperatures, and with the Rayleigh-Jeans formula for long wave-lengths and high temperatures.

Planck began by simply trying to find an equation which would fit the facts by reducing to the Wien equation for small values of λT, and to the Rayleigh-Jeans equation for large values of λT. As often in the history of physics he began by pure algebraical guesswork with no theoretical basis whatever. The theoretical basis followed the equation it seems to require.

Planck accepted the Rayleigh-Jeans expression for the number of degrees of freedom, but supposed that the average energy associated with each was not kT but

$$kT.\frac{x}{e^x-1}$$

where $x=\dfrac{h\nu}{kT}$, ν being the frequency of radiation of wavelength λ (so that $\lambda\nu=c$, the velocity of light) and h is an entirely new constant, later famous as Planck's constant.

Planck's equation thus became

$$E\lambda=8\pi kT\lambda^{-4}\frac{x}{e^x-1}.$$

The reader can satisfy himself that, since

$$e^x-1=x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots,$$

it follows that $\dfrac{x}{e^x-1}\to\dfrac{x}{x}\to 1$, higher powers of x becoming relatively negligible, as $\lambda T\to\infty$, since $x=\dfrac{h\nu}{kT}=\dfrac{hc}{k\lambda T}$, and $x\to 0$ therefore.

Planck's equation therefore approaches the Rayleigh-Jeans equation as λT approaches infinity.

As λT approaches zero, on the other hand, the difference between e^x and (e^x-1) becomes more and more negligible, and Planck's formula approaches Wien's. Thus if Planck's equation is correct, the Rayleigh-Jeans equation should tend to hold for long waves and high temperatures, and Wien's for short waves and low temperatures.

Planck's formula thus gave agreement with experiment; but its theoretical basis required assumptions which took him in the end much further than he wanted to go.

This dilemma of Planck's is profoundly interesting—or it should be—to any physicist who is not a mere technician.

In 1900 Planck described his work, destined to cause perhaps the greatest of all revolutions in human thought as yet achieved, as " only an interpolation formula found by happy guesswork." Any one who reads his two popular books, *Where is Science going?* and *The Philosophy of Physics* (Allen & Unwin, 1933 and 1936), will, I think, find that his general view was well expressed by the title of the first of them.

The physical justification of the formula seems then to have followed, not preceded, the formula itself. Planck thought it required that energy should be exchanged, given or received, not continuously but in definite amounts of a size fixed by the frequency of the radiation carrying the energy (not, as appeared in his formula, the wave-length, for wave-length is altered if radiant energy crosses the boundary between different transparent media, but frequency remains the same). Einstein it was who introduced the idea of the quantum of radiant energy travelling like a bullet through space, or like one of the particles required by Newton's Emission Theory of Light. Before long Einstein's Relativity Theory even gave these light-quanta definite mass in accordance with the $E = mc^2$ equation for the interchange of mass and energy; for by the quantum relation $E = h\nu$, so that $h\nu = mc^2$ and the equivalent mass of a quantum is given by $m = \frac{h\nu}{c^2}$. For visible yellow light of wave-length 6×10^{-5} cm. the mass of a quantum is thus given by

$$m = \frac{6{\cdot}6 \times 10^{-27} \times 3 \times 10^{10}}{(3 \times 10^{10})^2 \times 6 \times 10^{-5}} \doteq 3{\cdot}7 \times 10^{-33} \text{ gm.}$$

Let us return for a moment to Planck's equation

$$E_\lambda = 8\pi k T \lambda^{-4} \frac{x}{e^x - 1}$$

Where this differs mathematically from the corresponding (but wrong) equations of Rayleigh and Wien is that it is characteristic not of an integral but of a sum of terms of separate values.

In Planck's calculation these separate terms were possible energies for an oscillator, which could have energies 0, ϵ, 2ϵ, 3ϵ, 4ϵ, etc.—always a multiple of ϵ, a definite amount, or quantum, of energy. Planck therefore postulated the bold

idea that energy could only be transferred in definite units (like pouring peas from a pod, not water from a jug). His equation, which satisfied the facts of radiation, also required that the size of the quantum should be proportional to the frequency of the wave-length carrying its energy.

In any event, the Planck-Einstein development of the successful formula for black-body radiation finally led to the hypothesis that all radiant energy exists in space, or is held by or transferred from atoms or molecules, in definite amounts proportional to the frequency of the radiation or vibration concerned. If ϵ is the unit of energy concerned, then $\epsilon = h\nu$. where ν is the frequency and h a universal constant. The value of h accepted in 1949 was $6{\cdot}624 \pm 0{\cdot}002 \times 10^{-27}$ erg-seconds.

Writing the dimension equation for $\epsilon = h\nu$,

we have $$ML^2T^{-2} = h \times T^{-1}$$

so that $$[h] = [ML^2T^{-1}].$$

This is the dimension of energy $\times$ time (known as "action") or of momentum $\times$ length, or of angular momentum.

A reader who finds himself specially interested in this part of the subject would do well to read Chapter V of Richtmyer and Kennard's *Introduction to Modern Physics*, particularly section 83, which shows how Planck's formula, empirical at first, led him inevitably, and perhaps unwillingly, to a quantum theory of the transmission of energy.

Before Bohr applied the Quantum Theory to atoms, it was used successfully to show the way to the solution of two outstanding puzzles of physics—the photo-electric effect, and the falling-off of specific heats at low temperatures. Hertz had shown in 1887 that ultra-violet light falling on an uncharged conductor caused it to become positively charged. This was later shown to be due to the emission of electrons, which were somehow jerked out of the surface by the light. The velocity of emission of these electrons was measured, and was found to vary from zero up to a certain maximum, v, which (to the dumbfounding of all concerned) depended on the *frequency* of the light only. It was entirely independent of the intensity of the light, or the temperature of the surface, or the time for which the light had been falling. However faint the light, photo-electrons appeared at once. More

light produced more electrons, not faster electrons. On the wave-theory of light, this simply did not make sense. Einstein speedily explained these phenomena by means of the Quantum Theory. Suppose a quantum $h\nu$ of energy falls on a surface. Suppose an amount of work W is needed to get the electron out of the surface. Suppose all the energy of the quantum, which is left over after the work W has been supplied, is imparted to the electron. An electron of mass m will then acquire kinetic energy $\frac{1}{2}mv^2$.
Then

$$\tfrac{1}{2}mv^2 = h\nu - W.$$

Einstein introduced a new term, the "threshold frequency," the lowest frequency which will just get the electron over the threshold and out of the metal. He called this frequency ν_0. Obviously $W = h\nu_0$, and we may rewrite the equation

$$\tfrac{1}{2}mv^2 = h(\nu - \nu_0).$$

This equation gave a complete explanation of the facts so far as they concerned photo-electricity, but they caused a headache which has hardly yet subsided, at least in elementary laboratories, about interference. Both interference (detected photographically) and the photo-electric effect started immediately with exceedingly faint light, light so faint that it could only be carrying about one quantum per cubic metre. At least two overlapping quanta seemed necessary for interference. One highly concentrated quantum seemed necessary for each photo-electron. Yet both phenomena persisted unmistakably and obstinately.

It is from Einstein's conception of the quantum or bundle of energy, travelling as a single unit in space, that the modern term "photon" has been derived. It is now regarded as a kind of particle, having the velocity of light, but also having a definite amount of momentum (depending, as we shall show later, on its frequency). A photon may be defined as a quantum of radiant energy in space.

The specific heat problem is also much beyond the scope of this book, but a hint of the way the Quantum Theory bears on it can be given.

Dulong and Petit had proved theoretically that, on the basis of classical physics, the atomic heat (atomic weight × specific heat) of an element should be 3R where R is the gas

constant that occurs in the equation for a perfect gas, PV = RT. R has the value $8{\cdot}31 \times 10^7$ ergs, or $1{\cdot}98$ calories. The atomic heat should then by this theory be about 5·94 calories per degree (*e.g.* for copper at 20° C., $63{\cdot}6 \times 0{\cdot}091 = 5{\cdot}8$).

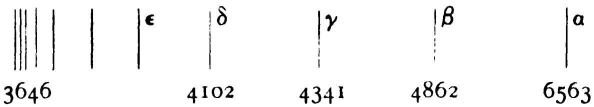
This law is not very far out for most elements at ordinary temperatures, but as the temperature approaches absolute zero the atomic heat also seems to approach zero. The problem seemed to have points of resemblance to the original radiation problem. In both a quantity associated with energy was found experimentally to approach zero when classical theory required that it should not. Einstein proposed that for the expression for atomic heat 3R should be replaced by $3R\frac{x^2e^x}{(e^x-1)^2}$, where $x = \frac{a}{T}$ so that the factor which appears squared in the denominator is precisely the factor which appeared in the denominator of Planck's Radiation Law, quoted above.

As the absolute temperature T approaches zero, $x \to \infty$, and the factor $\frac{x^2e^x}{(e^x-1)^2} \to 0$, so that the atomic heat $\to 0$. When T is large, x is small, and the whole factor $\frac{x^2e^x}{(e^x-1)^2}$ gets nearer and nearer to unity as T gets larger, so that the atomic heat approaches 3R. The present approach to the problem, though different from Einstein's, is on the same lines.

Bohr and the Spectrum of Hydrogen

The next great puzzle for classical theory was the series spectrum emitted by hydrogen. The Rutherford nuclear atom was accepted. The α-particle experiments gave a clear demonstration of its reality. For example, the Wilson cloud experiments could be made to show that α-particles usually travel their whole range in air in perfectly straight lines. Only very rarely did a deflecting collision happen, although the α-particle is much lighter than the oxygen or nitrogen nucleus. Geiger and Marsden found in 1909 that only about one α-particle in 5000 was diverted by collision when each particle passes about 1,000,000 atoms, and Rutherford in 1911

pointed out that this must mean that the cross-section of the nucleus is of the order of 2×10^{-10} of the cross-section of the whole atom. From other evidence it was known that the effective diameter of a whole atom is of the order of 10^{-8} cm., so that the diameter of the nucleus must be of the order of $10^{-8} \times \sqrt{2} \times 10^{-5}$, or about 10^{-13} cm. This being so, atoms must necessarily be almost completely empty structures, considerably emptier in proportion to size than the solar system. The hydrogen atom must therefore be a structure consisting of a single minute nucleus with a single electron describing an orbit about it. Yet this simple structure produced a most complicated and sharply defined spectrum. The Balmer series, which is characteristic of spectral series, gave lines running from 6563 Å to a limit at 3646 Å, as shown roughly below, the lines getting fainter and more crowded as they approach the limit.



Balmer discovered this series in 1885.

The first theory to bring order into spectral series was Rydberg's, in a paper published in the transactions of the Royal Swedish Academy in 1890.

He noticed that these lines were roughly connected by the equation

$$\lambda = \lambda_0 + \frac{C}{m + \mu}$$

where λ_0, C, and μ are constants and m is a series of integers. Then he tried wave-numbers instead of wave-lengths (*i.e.* number of waves in a centimetre) and found a very much better formula. Where n is the wave-number

$$n = n_0 - \frac{R}{(m + \mu)^2}.$$

In this n_0 is the wave-number of the limit, *i.e.* $\frac{1}{3646 \times 10^{-8}}$ cm.$^{-1}$ in the Balmer series. R was the famous Rydberg constant, which Rydberg got as 109,721·6 cm.$^{-1}$. It must not be confused with gas constant R (p. 542).

For the Balmer series Rydberg found $R = 4n_0$ and $\mu = 0$, so that the formula reduced to

$$n = \frac{1}{3646 \times 10^{-8}} - \frac{4}{3646 \times 10^{-8}} \times \frac{1}{m^2}$$

where m has integral values from 3 to ∞. Thus the first line, 6563 had a wave-number $\left(1 - \frac{4}{9}\right)$ of that of the limit, or in other words $6563 = \frac{9}{5} \times 3646$, which the sceptic will find to be true on test.

The next advance was made by Ritz in 1903. He noticed that the Balmer series could be represented by a much simpler formula

$$n = R\left(\frac{1}{2^2} - \frac{1}{m^2}\right)$$

where m can have all integral values from 3 upwards, n is the wave-number, and R is Rydberg's constant. The reader should test this on lines β, γ, δ, of the Balmer series, using the modern value of R.

The puzzle was to discover how in the world anything so simple as the Rutherford hydrogen atom could produce anything so complicated and so precise as the Balmer series (to say nothing of the other hydrogen series). It seemed obvious that the frequency of each emission line must be the frequency of revolution of the electron in its orbit under the influence of the electrostatic attraction of the nucleus. But if this were so, why should certain particular orbits alone be possible? Still worse, how could the atom radiate at all without the collapse of the electron into the nucleus, since a radiating electron—and a moving charge should radiate—ought to lose energy? What was the significance of Rydberg's constant R and why should the terms be spaced according to the $\frac{1}{2^2} - \frac{1}{m^2}$ rule? To all these questions no vestige of an answer suggested itself until Niels Bohr, a young research student working with Rutherford at Manchester, applied Planck's quantum to the problem in an entirely novel way. He began by applying to the Rutherford model of the hydrogen atom three quite extraordinary assumptions.

First Assumption

He assumed that only certain orbits are "permissible"; those in which the angular momentum of the electron is some multiple of $\frac{h}{2\pi}$, or, to put it in symbols,

$$mav = \frac{nh}{2\pi},$$

where m is the electron's mass, a the radius of its (circular) orbit, v its velocity, h Planck's constant, and n an integer.

No hint was given about the meaning of the word "permissible"; but one may recall here that a vibrating string or a bell may be said to have only certain modes of vibration permissible; in the string both the ends must be nodes.

(It is interesting to note the values of a which follow this assumption. Equating the radial force to the electrostatic attraction we get $\frac{mv^2}{a} = \frac{e^2}{a^2}$. Combining this equation with the first to eliminate v, we get $a = \frac{n^2h^2}{4\pi^2e^2m}$. Substituting the standard values for h, e, and m we get $a = 0{\cdot}53 \times 10^{-8}n^2$. Thus the diameter of the innermost permissible orbit for hydrogen would be $1{\cdot}06 \times 10^{-8}$ cm., and of the next $4{\cdot}24 \times 10^{-8}$ cm., quite sensible values.)

Second Assumption

The electron in its orbit does not radiate as it should on classical theory; but the atom as a whole does radiate when, and only when, the electron is transferred to some inner orbit of less energy; and when this happens and an amount of energy E is radiated, equal to the difference between the energies of the two orbits, then a single quantum of radiation of frequency ν is emitted, ν being determined by the quantum equation

$$E = h\nu.$$

Thus when an electron falls (losing energy) from an outer orbit of energy W_2 to an inner orbit of energy W_1, then the frequency ν is given by

$$h\nu = W_2 - W_1$$

The exact meaning of the "energy" of an orbit will be discussed in connection with the third assumption.

Third Assumption

When not radiating, the electron obeys the laws of classical physics in the most orthodox way.

This assumption enables one to give an exact meaning to the conception of the energy of an orbit. All energy of a system is referred to an arbitrary zero. A pound weight on a table 3 feet from the ground has 3 foot-pounds of energy if we assume that its zero of energy is when it is lying on the floor.

The energy of the electron in an orbit of radius a is thus the sum of its kinetic and potential energies, and the potential energy zero must be arbitrarily chosen. It is chosen to be when the electron is completely detached from the atom. Thus the potential energy in an orbit is negative and is given by

$$W_p = \int_{\infty}^{a} \frac{e^2}{r^2} dr = \left[- \frac{e^2}{r} \right]_{\infty}^{a} = - \frac{e^2}{a}.$$

Also

$$W_k = \frac{1}{2} mv^2$$

and the equation of circular motion gives us

$$\frac{mv^2}{a} = \frac{e^2}{a^2}.$$

so that

$$\frac{1}{2} mv^2 = \frac{1}{2} \frac{e^2}{a}.$$

Thus the total energy W is given by

$$W = W_p + W_k$$

$$\therefore W = - \frac{e^2}{a} + \frac{e^2}{2a} = - \frac{e^2}{2a}.$$

This ends the assumptions. Let us now collect the useful equations, and see whether we can find an expression for ν, the frequency of the emitted radiation, in terms we can measure. We have

$$mav = \frac{nh}{2\pi} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$h\nu = W_2 - W_1 = \frac{e^2}{2a_1} - \frac{e^2}{2a_2} \quad . \quad (2)$$

$$\frac{mv^2}{a} = \frac{e^2}{a^2} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Equation (1) really occurs twice, for a_1 with n_1 and for a_2 with n_2; and equation (3) occurs twice, for a_1 with v_1 and for a_2 with v_2.

We want to eliminate a_1, a_2, v_1, and v_2, and express $\lambda\left(=\frac{c}{\nu}\right)$ in terms of m, e, h, n_1, and n_2.

From (1) and (3) we get

$$v = \frac{mav^2}{mav} = \frac{e^2}{nh/2\pi} = \frac{2\pi e^2}{nh}. \quad . \quad (4)$$

From (2) and (3) we get

$$h\nu = \frac{e^2}{2a_1} - \frac{e^2}{2a_2} = \frac{1}{2}m(v_1^2 - v_2^2).$$

So from (4) $$h\nu = \frac{1}{2}m\left\{\left(\frac{2\pi e^2}{n_1 h}\right)^2 - \left(\frac{2\pi e^2}{n_2 h}\right)\right\}$$

hence $$\frac{hc}{\lambda} = \frac{2\pi^2 me^4}{h^2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right),$$

so $$\frac{1}{\lambda} = \frac{2\pi^2 me^4}{h^3 c}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

One further correction is necessary. We have assumed that the mass m of the electron is negligible compared with the mass M_H of the hydrogen nucleus. Really this is not so, for m and M_H both revolve round their common centre of gravity. It can be shown fairly easily—the reader should do it as an exercise in dynamics—that we can correct for this discrepancy by putting $\frac{mM_H}{m + M_H}$ instead of m. We thus get

finally $$\frac{1}{\lambda} = \frac{2\pi^2 mM_H e^4}{h^3 c(M_H + m)}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

This is the miracle-equation, as it must have seemed at the time. It precisely follows Ritz's form for the Balmer series, and by substituting the known values for m, M_H, e, h, and c, we get the original Rydberg's constant outside the bracket. The reader should do this for himself. If it be asked how Bohr obtained a value for h, we can reply that quite a number of values were already available by 1913—from radiation, the photo-electric effect, and the theory of specific heat, to begin with.

Here was conclusive, finally conclusive, evidence that Bohr's irrational explanation of Rydberg's constant contained, though probably in a disguised form, the truth. It was published in 1913. On the title-page of Sir Charles Darwin's book, *The New Conceptions of Matter* (Bell, 1931), is the inscription

DEDICATED AS A TRIBUTE OF
ADMIRATION AND AFFECTION
TO
NIELS BOHR
THE ACKNOWLEDGED LEADER
AMONG THOSE WHO HAVE SHOWN
HOW THE SECRETS OF NATURAL PHILOSOPHY
MAY BE UNDERSTOOD

Now, why did Sir Charles Darwin write so of Bohr in the lifetime of Rutherford and Einstein and Planck? The view he expresses is not meant to be a personal one, for he speaks of " the acknowledged leader "; and he is indeed expressing a feeling which is still, in 1949, pretty general among atomic physicists. Nor is he diminishing the stature of Einstein, Planck, or Rutherford, each of whom is in a different way the acknowledged leader. Bohr's pre-eminence is of a special kind; and, to understand what this kind is, let us consider one of the transitions from orbit to orbit during which, or at the instant of which (on Bohr's theory) the atom radiates.

How long does the transition take? What makes it happen? Where is the electron after it has left one orbit but before it has entered the next? Why are only certain orbits permissible? Who, or what, permits them? Why are the classical laws neither obeyed universally nor disobeyed universally, but obeyed selectively?

To these questions I suggest the answer that Bohr, whether consciously or unconsciously, was not regarding time and space as inflexible and superhuman, but as useful adjuncts to our attempts to explain the world of our experience up to the point we can reach in that explanation with due humility. What, I think, made Bohr the " acknowledged leader " was the thorough-going boldness of his use of what lies beyond the limits of our visual imagination. It is in this that he has as his followers both the schools of thought in atomic theory

which succeeded in some of the problems in which Bohr's first method of attack failed; the quantum mechanics of Heisenberg, Born, and Jordan, and the wave mechanics of De Broglie, Schroedinger, and Dirac; for the work of both these schools would have been inconceivable if Bohr had not taken the first big step.

We can at this point clarify a little some of the difficulties of Bohr's elementary theory. If time and space are useful concepts rather than inescapable absolutes—the philosopher Kant called space " the form of all appearances of outer sense " and time " the form of inner sense "—it is not surprising that our conception of them works for things of an order of size not too remote from ourselves—say things bigger than a molecule but smaller than the solar system. The Relativity Theory suggests divergencies from our conceptions in the very great (and inside the atom the very swift-moving) and the Quantum Theory in the very small.

There is an interesting point of temperament here. At Cambridge in 1922–3, as a graduate student, I heard Eddington, Rutherford, and Bohr. It struck me that Eddington would have preferred something mystical, or at least odd, in a theory. Rutherford would have avoided it with determination in one of his own theories. But Bohr would have had no preferences whatever. His mind struck me as completely open. It was typical of him that the electrons in his atom described strictly classical orbits, but did not radiate as good classical moving charges should have. He seemed outstandingly what a theoretical scientist should be, a man willing to try anything in a very odd world in which we should regard ourselves as inexperienced newcomers.

Bohr's elementary theory, when adapted by Sommerfeld to work for elliptical as well as circular orbits, for a spinning electron, for relativity corrections associated with very high speeds, and for some rather curious restrictive rules, was found to work with complete correctness for any atom with only one electron. It has been proved to work with singly-ionized helium, doubly-ionized lithium, triply-ionized beryllium, fourfold-ionized boron, fivefold-ionized carbon, and sixfold-ionized nitrogen, produced by a remarkable new technique of specially hot sparks.

It also worked for the characteristic X-rays of the elements,

and so provided a theory for Moseley's fundamental experiments which stood the periodic table on really secure legs for the first time.

The Rydberg constant, however, had a slightly different value for each element, since it always contains as a multiple the factor $\frac{M}{M + m}$, and this gets very slightly larger as the element chosen gets heavier. For an element other than hydrogen, carrying a nuclear charge Ze where Z is the atomic number, if we put $\mu = \frac{mM_Z}{m + M_Z}$, we get for the Bohr equation

$$\frac{1}{\lambda} = \frac{2\pi^2 \mu Z^2 e^4}{h^3 c}\left(\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right).$$

(The Z^2 appears because W becomes $-\frac{Ze^2}{2a}$ instead of $-\frac{e^2}{2a}$ (see p. 551) when the nuclear charge becomes Ze instead of e). This should be easily seen by a reader who can follow the argument of p. 551.)

The hydrogen equation also explained the ultra-violet Lyman series and the infra-red Paschen and Brackett series. We have seen that for the Balmer series $n_1 = 2$. For the Lyman series $n_1 = 1$. For the Paschen, $n_1 = 3$. For the Brackett, $n_1 = 4$.

We have already seen the values of the radius a of the electron for the permissible orbits. The frequency of the electron in its orbit is also interesting, not only because we can see that in general it is different from the frequency of the emitted light, but also because the investigation of it leads us to another very odd fact.

The period of revolution $T = \frac{2\pi a}{v}$ obviously.

The frequency is then $\frac{1}{T}$, or $\frac{v}{2\pi a}$.

We have already seen that $a = \frac{n^2 h^2}{4\pi^2 m e^2}$,

and that $$v = \frac{2\pi e^2}{nh}$$

So we have $$\frac{1}{T} = \frac{1}{2\pi} \cdot \frac{2\pi e^2}{nh} \cdot \frac{4\pi^2 me^2}{n^2h^2} = \frac{4\pi^2 me^4}{n^3h^3}.$$

The frequency of the emitted light, on the other hand, is given by
$$\frac{c}{\lambda} = \frac{2\pi^2 me^4}{h^3}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

Let us consider the ratio of these frequencies to each other. The factor $\frac{2\pi^2 me^4}{h^3}$ cancels, and we have left

$$\frac{2/n^3}{1/n_1^2 - 1/n_2^2}$$
$$= \frac{2n_1^2 n_2^2}{n^3(n_2^2 - n_1^2)}$$
$$= \frac{2n_1^2 n_2^2}{n^3(n_2 + n_1)(n_2 - n_1)}.$$

Now consider what happens for a transition from one very high orbit to the next below it. Put $n = n_1 = (n_2 - 1)$.

Then we may put $2n$ for $(n_2 + n_1)$, and 1 for $(n_2 - n_1)$, and we get

$$= \frac{2n^4}{n^3 \times 2n \times 1} = 1$$

The frequencies are the same; or to be more exact, the higher the quantum number the more nearly does the frequency emitted for a transition from an orbit to the next below approach the frequency of the electron in its orbit; the very frequency the atom ought to emit on the classical theory.

This illustrates one of the most important principles enunciated by Bohr, the Correspondence Principle, which states that the higher the quantum number involved the more nearly does the state of affairs approach to what would be normal on classical theory. Compare the success of Rayleigh's formula for radiation at one end of the spectrum but not the other, and the success of classical specific heat theory at high temperatures but not at low. Quantum Theory repeatedly gives the impression of something which looks very different when seen under a microscope, like a drop of water which contains a lot of things which give no hint of their presence when viewed macroscopically. Bohr's simple theory, though not really satisfactory for an atom with more than one planetary electron (since the three-body problem cannot be

solved analytically) did lead to the more powerful though less intelligible methods of quantum mechanics and wave mechanics. It also provided a method of considering the physical and chemical properties of all the elements, and Bohr's own account of this part of his theory is given in a 1922 reprint of three of his lectures, *The Theory of Spectra and Atomic Constitution* (Cambridge University Press). Typically, he discusses the characteristics of atoms with Atomic Numbers up to 118 (the presumed Atomic Number of the next inert gas after radon), mildly regretting (as I heard him regret in a lecture at Cambridge in 1922) that the atoms with Atomic Numbers above 92 had not succeeded in existing, or perhaps in continuing to exist; but the recently found (or recently synthesized) atoms with Atomic Numbers 93 to 96 seem less unstable than Bohr then feared, so perhaps his 1922 longing for at least 26 more elements may still be satisfied.

After this note on the correspondence principle, it is appropriate to give a note on the other principles by which Bohr developed his theory, though to follow it any further would be beyond the scope of the book.

Up to now we have dealt only with circular orbits for a single electron. It is, however, necessary to deal with many electrons in elliptic orbits of varying eccentricity, whose major axes are considered to revolve round the nucleus because the relativity change of mass of the electron as it goes round causes the orbit to precess, and the electron to trace out a kind of rosette. The electron is also assumed to have one of two kinds of spin. To cope with this situation it has been found necessary to have four quantum numbers; n, the principal quantum number giving the total energy; l, the subordinate quantum number, which may have any value from 0 to $(n - 1)$; m, the magnetic quantum number, which may have any value from $+ l$ to $- l$; and σ, the spin quantum number, which (oddly enough) can have values either $\frac{1}{2}$ or $-\frac{1}{2}$.[1]

What brought order into this system, and linked it on to the periodic system, was Pauli's principle, that no 2 electrons could ever have all 4 quantum numbers the same. Applying Pauli's principle, and remembering that electrons at the

[1] The 1949 arrangement of the 4 quantum numbers is rather different, but more difficult to begin with.

K-level are 1-quantum electrons, at the L-level 2-quantum, and so on, we can see what happens from the table below. (For an account of these levels, see page 572.)

From this we can see that on Pauli's principle, allowing for the fact that there are always two possible values for the spin quantum number, there can be no more than 2 electrons in the K-ring, 8 in the L-ring, 18 in the M-ring, and 32 in the N-ring.

The numbers 2, 8, 18, 32 are the numbers of the elements in the various periods of the periodic table; and beyond this hint we cannot go in this book.

$n=1$ (K-ring)	$l=0$	$m=0$	$2\times 1=2$
$n=2$ (L-ring)	$l=0$	$m=0$	$2\times 4=8$
$n=2$,,	$l=1$	$m=-1, 0, +1$	
$n=3$ (M-ring)	$l=0$	$m=0$	
$n=3$,,	$l=1$	$m=-1, 0, +1$	$2\times 9=18$
$n=3$,,	$l=2$	$m=-2, -1, 0, +1, +2$	
$n=4$ (N-ring)	$l=0$	$m=0$	
$n=4$,,	$l=1$	$m=-1, 0, +1$	
$n=4$,,	$l=2$	$m=-2, -1, 0, +1, +2$	$2\times 16=32$
$n=4$,,	$l=3$	$m=-3, -2, -1, 0, +1, +2, +3$	

As a magnetic quantum number has been mentioned, we can point out here that the Bohr atom can be seen to be associated with at least three kinds of rotating or revolving charge. There are electrons in orbits, spinning electrons, and spinning nuclei. Magnetic effects arise from revolving or rotating charges, and we would therefore expect to find that all magnetic phenomena originate in the atom; but to follow this clue here would take us too far.[1]

The Beginnings of Wave-mechanics

The following short account may give the beginner a hint of the method by which Schroedinger, using an astonishing idea of de Broglie's, and applying it to the general differential equation of wave-motion in an elastic medium, arrived (without Bohr's arbitrary assumptions) at Bohr's results for a one-electron atom.

[1] Interested students should try H. E. White's *Introduction to Atomic Spectra,* to which J. C. Speakman's *Modern Atomic Theory,* Chapters VI and VII, is an excellent stepping-stone.

De Broglie (pronounced Debroy) published his first paper on wave-mechanics in 1925. He began by saying that we have discovered that we fail to interpret experience unless we regard radiant energy in a particle aspect as well as a wave aspect, and that we may therefore regard these two aspects of reality as in some way complementary. If we do this, we may reasonably expect that material particles should also be regarded in these two complementary aspects, and hence that a material particle, such as an electron or proton, may be associated with a wave-length. His emphasis was entirely on the wave-length, not at all on the medium. He proceeded to argue by analogy to the probable value of the wave-length, in order to make an experimental test possible, for in physics no speculation is of any value unless it can be tested experimentally.

Several ways of thinking lead us by analogy to de Broglie's equation on which Schroedinger based his development of wave-mechanics; and of these ways we will here give the two easiest. Both lead directly to the equation

$$\lambda = \frac{h}{mv},$$

in which h is Planck's constant and mv is the momentum of the particle.

First Analogy. A stretched string can vibrate in various modes. It can have one loop, two loops, three loops, and so on, but only an integral number of loops. It can only emit frequencies associated with an integral number of loops, not intermediate frequencies. The system has stationary states associated with whole numbers, a definite frequency being associated with each state and its corresponding whole number. In this a stretched string is analogous to a Bohr atom. As each Bohr orbit is associated with a particle, the orbital electron, and a length, the perimeter of the orbit, let us try the effect of supposing that the perimeter of the one-quantum circular orbit of a Bohr hydrogen atom is the wave-length associated with an electron moving with the Bohr orbital velocity.

We have, by Bohr's first assumption, already discussed in this chapter,

$$mav = \frac{h}{2\pi},$$

from which

$$2\pi a = \frac{h}{mv}.$$

But $2\pi a$ is the perimeter of the orbit, and hence by analogy we at once get

$$\lambda = \frac{h}{mv}.$$

Second Analogy. A. H. Compton's experiments on the scattering of X-ray photons by electrons showed that the photon must be regarded as having momentum p such that

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}.$$

De Broglie argued that a similar relation might be expected between the momentum mv of a massive particle and the wave-length λ associated with it. We then have

$$mv = \frac{h}{\lambda}$$

and hence $\lambda = \dfrac{h}{mv}$ as before.

Third Analogy. The same result can also be obtained by supposing the particle to be associated with a group of waves in a dispersive medium, the group velocity v being the particle velocity, and the wave-length that associated with the particle wave-length. Let the wave velocity be u.

The argument is longer and more difficult than the two we have given (see Richtmyer and Kennard's *Introduction to Modern Physics*, section 107), but it leads to the same equation, and also to the surprising conclusion that $uv = c^2$; so that, since v is always less than c for a material particle, u must always be greater than c, and so, by the first postulate of the Relativity Theory, cannot represent the passage of anything real.

It should be noted that in the equation

$$\lambda = \frac{h}{mv}$$

m is not the rest-mass m_0 of the electron, but the mass proper to the velocity according to the relativity relation

$$m = \frac{m_0 c}{\sqrt{c^2 - v^2}}.$$

In calculating the wave-length of a particle one has to decide where to begin to allow for the relativity increase of mass

with velocity. A simple calculation shows that c is 1% greater than $\sqrt{c^2 - v^2}$ when $v = \frac{1}{7}c$ or about $4{\cdot}3 \times 10^9$ cm./sec., and that an electron with this velocity has fallen through about 5,000 volts. When we need not allow for the increase of mass with velocity we have, for an electron falling through a potential difference of V absolute units,

$$\mathrm{V}e = \tfrac{1}{2}m_0v^2$$

$$\therefore\quad 2\mathrm{V}m_0e = (m_0v)^2$$

$$\therefore\quad m_0v = \sqrt{2\mathrm{V}m_0e}$$

$$\therefore\quad \lambda = \frac{h}{\sqrt{2\mathrm{V}m_0e}}.$$

For a P.D. of 5,000 volts $\mathrm{V} = \frac{5000}{300}$ absolute units, and substituting the standard values for h, m_0, and e we get $\lambda = 0{\cdot}17$ Å.

The equations we already have tell us that, for a given value of m, λ varies inversely with the velocity of the particle, or inversely as $\sqrt{\mathrm{V}}$; and, for a given value of v or V, inversely as the mass of the particle. Thus a 100-volt electron has a wave-length of about 1·2 Å, a 100-volt α-particle a wave-length of 0·015 A°, an α-particle from radium with a velocity of $1{\cdot}5 \times 10^9$ cm./sec. has $\lambda = 6{\cdot}6 \times 10^{-5}$ Å. For a β-particle moving at about 0·9 c, $\lambda =$ 80·0224 Å, the relativity correction reducing λ by about 17 %.

These numerical results were experimentally confirmed by Davisson and Germer for slow electrons, with energies up to 370 ev, in 1927 and 1928, and by G. P. Thomson in 1931 for fast electrons. By 1938 the de Broglie equation had been verified for electron energies up to 64 kilovolts. It has also been verified for α-particles, protons, and monochromatic beams of various molecules.

We now turn to Schroedinger's application of de Broglie's equation to the general wave-equation

$$\frac{d^2\phi}{dx^2} = \frac{1}{c^2}\frac{d^2\phi}{dt^2}$$

in which ϕ is the displacement at time t and distance x along the x-axis of wave-motion with velocity c.

Suppose this wave-motion has wave-length λ, we can solve

the equation by introducing a new variable ψ which is a function of x, but not of t, so that

$$\phi = \psi e^{\frac{2\pi i c t}{\lambda}}$$

then

$$\frac{d\phi}{dt} = \frac{2\pi i c}{\lambda}\psi e^{\frac{2\pi i c t}{\lambda}}$$

$$\frac{d^2\phi}{dt^2} = -\frac{4\pi^2 c^2}{\lambda^2}\psi e^{\frac{2\pi i c t}{\lambda}}$$

On the other hand, performing a partial differentiation of ϕ with respect to x we get

$$\frac{d\phi}{dx} = \frac{d\psi}{dx} e^{\frac{2\pi i c t}{\lambda}}$$

$$\frac{d^2\phi}{dx^2} = \frac{d^2\psi}{dx^2} e^{\frac{2\pi i c t}{\lambda}}.$$

Now we can substitute for $\frac{d^2\phi}{dx^2}$ and $\frac{d^2\phi}{dt^2}$ in the general wave-motion equation we started with, and we get

$$\frac{d^2\psi}{dx^2}\cdot e^{\frac{2\pi i c t}{\lambda}} = \frac{1}{c^2}\left(-\frac{4\pi^2 c^2}{\lambda^2}\right)\psi\cdot e^{\frac{2\pi i c t}{\lambda}}$$

from which, dividing through by $e^{\frac{2\pi i c t}{\lambda}}$ cancelling c^2 and bringing the right-hand side across, we get

$$\frac{d^2\psi}{dx^2} + \frac{4\pi^2}{\lambda^2}\psi = 0.$$

Now, we substitute de Broglie's equation for λ, $\lambda = \frac{h}{mv}$, and get

$$\frac{d^2\psi}{dx^2} + \frac{4\pi^2 m^2 v^2}{h^2}\psi = 0$$

in which ψ is the amplitude of the wave-motion associated with the moving particle.

We now get rid of v^2 by expressing it as a difference between total energy and potential energy, since we can quantize the total energy and express the potential energy of an electron in terms of charge and distance from the nucleus. (We should understand that this procedure was not by any

means the result of inevitable mathematical logic. It was again successful guesswork by analogy, justified by experimental verification of results proceeding from it. How many unsuccessful guesses were made?)

Now *Total Energy* W = *Potential Energy* V + *Kinetic Energy* $\frac{1}{2}mv^2$

$$\therefore \quad W - V = \tfrac{1}{2}mv^2$$

$$v^2 = 2\left(\frac{W - V}{m}\right).$$

Hence by substitution we get

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}(W - V)\psi = 0$$

and if we generalize this for all three axes we get

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(W - V)\psi = 0$$

$$\left(\text{where } \nabla^2\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right)$$

At this point Schroedinger applied this general equation to the particular case of the one-electron atom with an electron of charge e revolving round a nucleus of charge Ze in an orbit of radius r.

For such an atom the potential energy V of the electron is given by

$$V = \int_0^r \frac{Ze^2}{r^2}\,dr = -\frac{Ze^2}{r}$$

by the ordinary coulomb law. Thus the equation becomes

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}\left(W + \frac{Ze^2}{r}\right)\psi = 0$$

and this is the famous Schroedinger wave-equation.

Its solutions must satisfy boundary conditions so that ψ and its derivative must be everywhere single-valued, continuous, and finite. These solutions are called eigen functions or eigen values. The dictionary meanings of the word *eigen* are " own, proper, peculiar, special, strange, queer, difficult, ticklish, delicate, and specific," the last being the physicist's

meaning, but all of them being somehow instructive. It is actually found that eigen values for W require that n (see p. 550) should be a whole number, 1, 2, 3, 4, . . . , and these values of W are identical with those obtained by the simple Bohr theory, as already done in this book. The reason why these integral values of n are inevitable is beyond the scope of this book, but it may perhaps be helpful to suggest that only for these values do the de Broglie waves in the electric field of the nucleus reinforce one another; otherwise they are mutually destructive in rather the way that light-waves are mutually destructive in the dark parts between bright fringes caused by diffraction or interference.

Critical Potentials in Gases

In 1914, the year after the appearance of Bohr's first paper on the hydrogen spectrum, Franck and Hertz published a

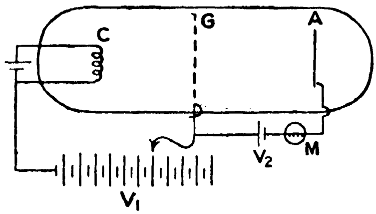


FIG. 267

paper on some work which confirmed Bohr's results in a striking, and at that time unexpected, way. This work, on the ionization and resonance potentials in gases, was the beginning of a long and important series of experiments from which many fundamental facts, about energy levels in many elements, have emerged.

The general idea of these experiments is shown in Fig. 267. Electrons from a cathode C, in a nearly exhausted glass tube, are accelerated by a potential V_1 towards an anode A. On the way they pass through a grid G; and—here is the special point of the arrangement—a *reverse* voltage V_2 of very small value compared with that of V_1, probably about half a volt, is applied between A and G, so that any electrons between A and G, without enough energy to carry them on to A against

this small field V_2, get carried back to G. There is thus a small current between A and G which is recorded by the sensitive galvanometer M. Only a trace of the gas under test is let into the tube, and its pressure is accurately adjusted in relation to the dimensions of the tube and the distances between the electrodes to ensure that most of the collisions between the accelerated electrons and the gas molecules will happen between G and A. The current flowing between G and A was plotted against gradually increasing values of V_1.

This current did not vary with V_1 in at all a uniform manner. Instead, it showed a succession of peaks corresponding to particular voltages for particular gases. Voltages showing peaks of current were obviously voltages which gave the electron an amount of energy which it could lose all at once in an inelastic collision. It was then left between G and P with no energy, and drifted back to G under the influence of the field V_2. Apparently the electrons did not always lose some energy; instead they lost either all or none.

This result, surprising at the time, fitted Bohr's atomic theory based on the Quantum Theory, and actually the observed results fitted the calculated values beautifully. Each inelastic collision occurred when the electron had an energy whose quantum gave a wave-length emitted by the atom. If V is the value giving a peak, then

$$Ve = \tfrac{1}{2}mv^2 = h\nu$$

where ν is the frequency of a known line.

Hydrogen, for example, showed peaks at 10·15 volts, 12·05, 12·70, several more, and finally, the largest, 13·54. The voltages calculated from observed hydrogen lines were 10·17, 12·05, 12·71, . . . 13·56.

The last one corresponded to complete ionization of the H-atom. Each of the others was called a resonance potential and occurred when the atom got enough energy to cause a particular shift between two energy-levels, without ionizing the atom. Helium showed resonance for 19·75, 20·55, 21·2, 22·9, 23·6, etc., volts; and ionization for 24·5 volts. Helium gives the highest values of any element. Mercury gives 4·68, 4·9, 5·47, etc., with ionization at 10·38. Some atoms have a second or third ionization potential, corresponding to energies

at which two or three electrons are knocked out. With normally diatomic gases like hydrogen there was confusion about the binding energy of the H_2 molecule, but the effect was confined to atoms by running the apparatus at not less than 2800° K. at which H_2 is mostly dissociated by thermal effects

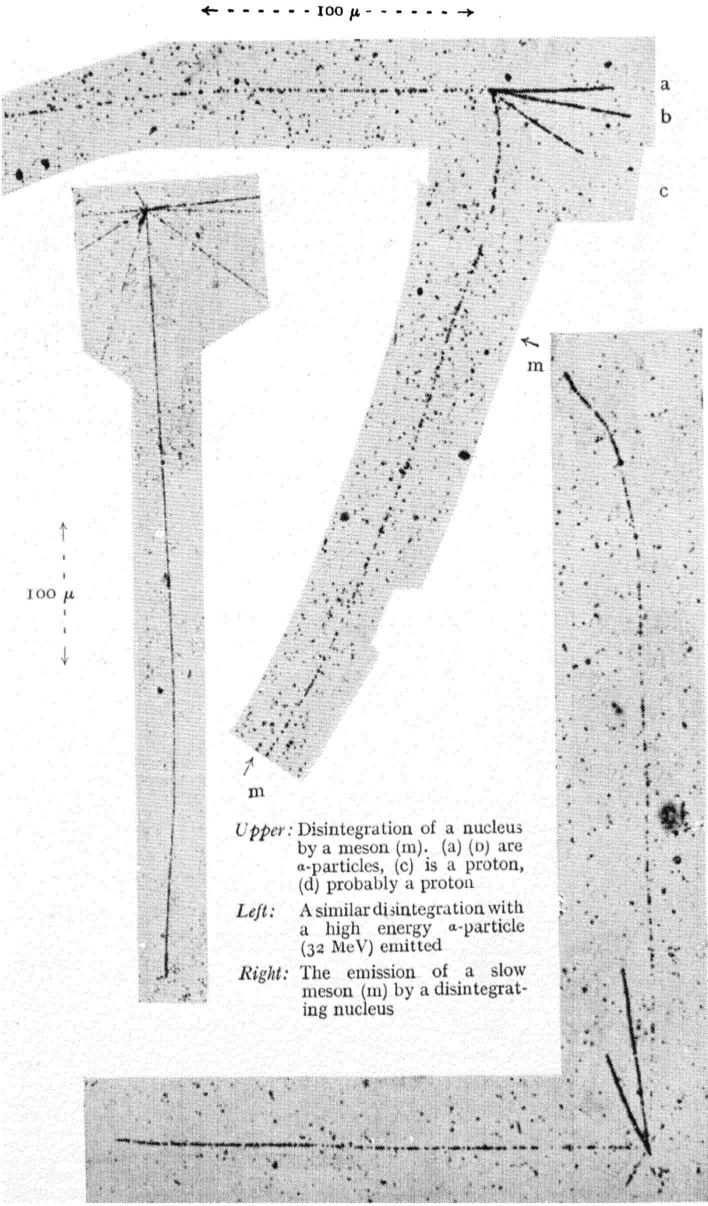
The ionization potentials of the inert gases get less as the atomic number increases; thus He, Ne, A, Kr, Xe have respective I.P.s of 24·5, 21·5, 15·7, 13·3, 11·5 volts.

The whole of this work gives a most fascinating picture of the Quantum Theory in action, showing, as it directly does, that an atom can take in (and give out) energy in discrete jumps only.

NOTE ON PLATES

These plates and those facing pages 590 and 591 are reproduced by permission of Prof. C. F. Powell, F.R.S., from Powell and Occhialini's *Nuclear Physics in Photographs*, and illustrate the authors' photo-micrograph technique developed at the Wills Physical Laboratory, Bristol University, since 1945. The events they illustrate are discussed in chap. vi, Part III.

Plate I



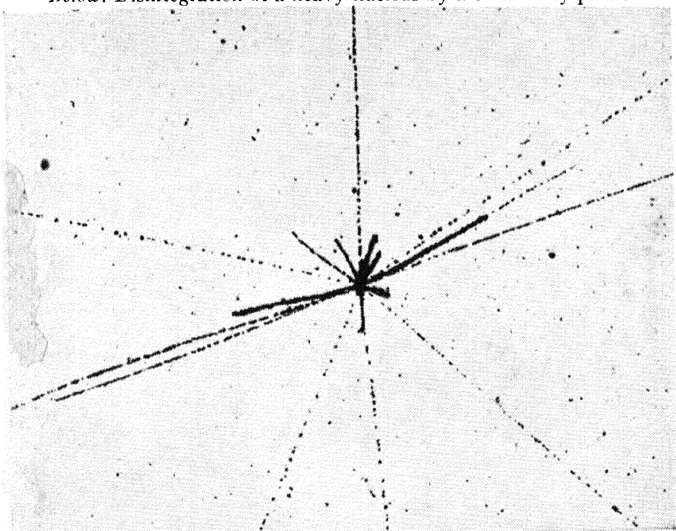
Upper: Disintegration of a nucleus by a meson (m). (a) (b) are α-particles, (c) is a proton, (d) probably a proton

Left: A similar disintegration with a high energy α-particle (32 MeV) emitted

Right: The emission of a slow meson (m) by a disintegrating nucleus

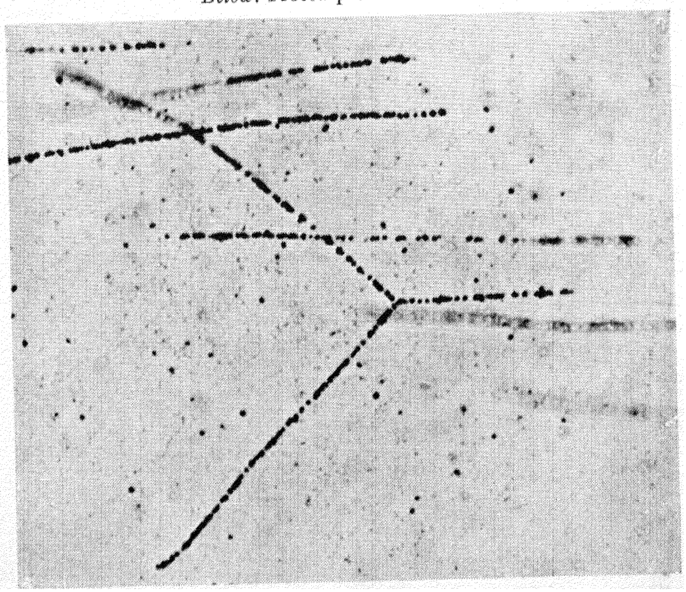
Plate 2

Below: Disintegration of a heavy nucleus by a cosmic ray particle



← - - - - - 100 μ - - - - - →

Below: Proton-proton collision



Line-spectrum of X-rays

Curves like those shown in Fig. 269 in the last section happen for a particular collection of target (or anode) materials up to only a certain value of V. Up to this value the shape of the curve is completely independent of the target material. But for each kind of target material there is a value for V above which the curve begins to get humps on it, as in Fig. 270. These humps do not appear gradually. There is a certain threshold value of V above which they suddenly appear, and this fact is quite easy to understand.

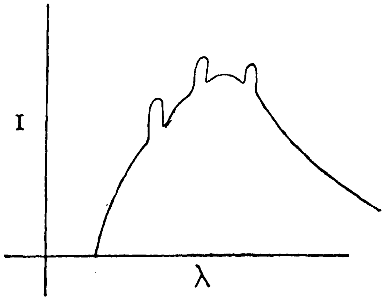


FIG. 270.

When V is just enough to kick an electron out of one of the lower energy-levels of an atom of the target, the atom emits its own characteristic X-rays, with definite frequencies characteristic of itself, as a result of the rearrangements in the electronic structure which must follow the expulsion of one of its electrons. The origin of line-spectra of X-rays is thus utterly different from the origin of continuous-spectrum X-rays.

If we follow on with Bohr's theory of the hydrogen spectrum, as given at the end of Chapter III, the whole business of line-spectrum X-rays becomes quite simple to understand because it all happens naturally and makes sense, and we shall therefore develop the subject logically instead of historically in this part.

At the end of Chapter III we had proved that for an atom with an electronic charge of Ze and a mass of M_2, but so much ionized that only one planetary electron is left, the wave-length for any line of an emission series is given by

$$\frac{1}{\lambda} = \frac{2\pi^2\mu Z^2e^4}{h^3c}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

where

$$\mu = \frac{mM_2}{m + M_2}.$$

If by R we mean R_∞, the value of Rydberg's constant which

would obtain if we could regard μ as for practical purposes equal to m, then

$$R = \frac{2\pi^2 me^4}{h^3 c} = 1\cdot097 \times 10^5 \text{ cm.}^{-1}.$$ [1]

We thus have, shortly, for hydrogen where $Z = 1$,

$$\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

For the Lyman series, we have $n_1 = 1$, $n_2 = 2, 3, 4, \ldots$

For the Balmer series, we have $n_1 = 2$, $n_2 = 3, 4, 5, \ldots$

For the Paschen series, we have $n_1 = 3$, $n_2 = 4, 5, 6, \ldots$

For the Brackett series, we have $n_1 = 4$, $n_2 = 5, 6, 7, \ldots$

For each series the lines are emitted when an electron falls back from any one of a number of outer energy-levels to a particular inner one which is characteristic of the series.

For the Lyman series the electron always falls back to the 1-quantum level (radius of circular orbit $0\cdot527 \times 10^{-8}$ cm., see p. 573). Let us call this energy-level the K-ring.

For the Balmer series the electron always falls back to the 2-quantum level (radius of circular orbit $2\cdot108 \times 10^{-8}$ cm., 4 times as large, or 2^2 times as large as the K-ring orbit, see p. 573). Let us call this energy-level the L-ring.

For the Paschen series the electron always falls back to the 3-quantum level (radius of circular orbit 3^2 times as large as the K-ring orbit). Let us call this energy-level the M-ring.

For the Brackett series the electron always falls back to the 4-qauntum level (radius of circular orbit 4^2 times as large as the K-ring orbit). Let us call this energy-level the N-ring.

Let us now represent the four hydrogen series, or rather the transitions which cause them both, semi-pictorially with the electron orbits associated with the various energy-levels shown roughly to scale (assuming that the radii we have calculated so carefully have a real meaning). This is in Fig. 271. Then we will represent them diagrammatically, not in any way to scale, merely showing energy-levels. We will show the energy-levels as far out as the 6-quantum level, radius 36 times as great as the K-ring radius.

Let us now in Fig. 272 give a diagrammatic representation of these energy-levels. This time, instead of spacing the levels in

[1] The 1949 best value of R_∞ is $109737\cdot30 \pm 0\cdot017$ cm.$^{-1}$.

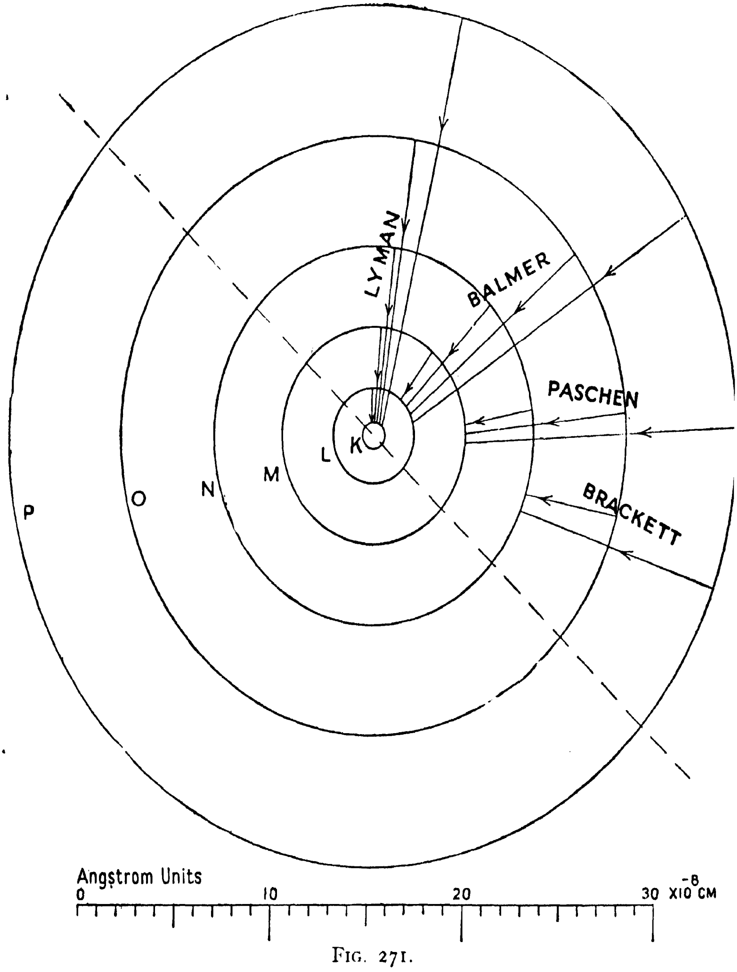


Fig. 271.

proportion to the calculated diameters of the circular orbits, let us space them in proportion to the energy involved in a transition from one to another. (The absolute energy of the levels has no meaning, as it must be measured from an arbitrary level; and zero energy is taken as the energy of the electron removed from the influence of the nucleus, so that the energy of each level is negative, the K level having the largest negative energy.)

We see from the work already done that the energy-levels are given by

$$h\nu = \frac{hc}{\lambda} = \frac{2\pi^2\mu Z^2 e^4}{h^2} \cdot \frac{1}{n^2},$$

where n is 1 for K, 2 for L, 3 for M, and so on.

If for hydrogen we put $Z = 1$ and $R = \frac{2\pi^2 me^4}{h^2 c}$, as before, we have

$$E = h\nu = \frac{hcR}{n^2} \text{ for } n = 1, 2, 3, 4, 5, 6.$$

Now $\quad hcR \doteq 6{\cdot}624 \times 10^{-27} \times 3 \times 10^{10} \times 1{\cdot}097 \times 10^5$

$\doteq 2{\cdot}17 \times 10^{-11}$ ergs $\doteq 217 \times 10^{-13}$ ergs.

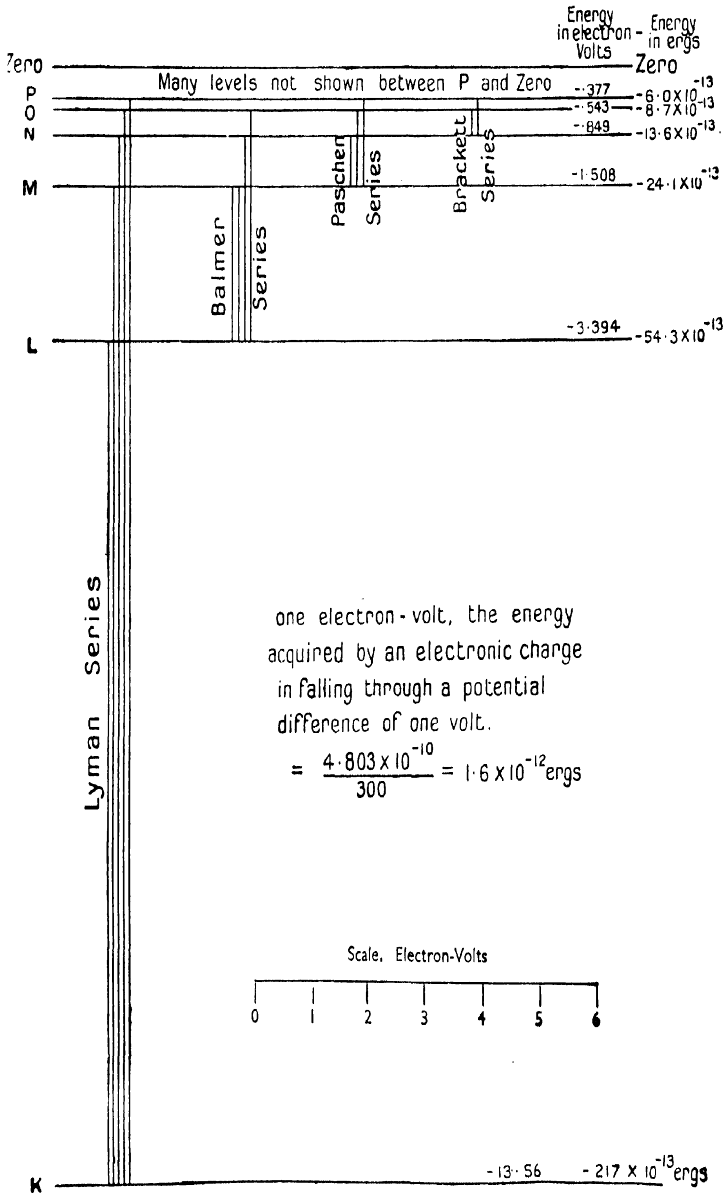
So the K level will have this (negative) energy.

The L, M, N, O, P levels will have $\frac{1}{4}$, $\frac{1}{9}$, $\frac{1}{16}$, $\frac{1}{25}$, $\frac{1}{36}$ as much, and have therefore, near enough for our diagram, energies of 54·3, 24·1, 13·6, 8·7, 6·0 $\times 10^{-13}$ ergs.

A diagrammatic scheme for energy-levels and transitions for circular orbits in hydrogen is given in Fig. 272.

These energies are also shown in electron-volts. The reader should note the calculated K-energy of −13·56 volts and compare it with the voltage required to ionize the H-atom (see p. 565).

These energy-levels are universal for all elements. The six lowest are always known as K, L, M, N, O, P, working outwards from the nucleus. The K level is always single, but all the others, for reasons beyond the scope of this book, are multiple (but see p. 578). Radiation due to transition to one of them from another is called X-radiation simply when its energy is high enough (or its wave-length is short enough, which is, of course, the same thing). The first characteristic radiation which counts as X-rays is the K-radiation for sodium, the first element of the third group in the periodic table, for which $Z = 11$. For each atom there is an X-ray series based on each level that is deep enough to give X-rays. Thus all atoms from ${}_{11}$Na to ${}_{92}$U give K series. L series begin with ${}_{29}$Cu, M series with ${}_{66}$Dy, and N series with ${}_{83}$Bi. Levels which occur in a series as helping to produce transitions, but not as basic levels, show up more, though levels containing regular sets of electrons in normal neutral atoms do not go



Energy in electron-Volts
Energy in ergs
Zero
Zero
Many levels not shown between P and Zero
P
O
N
M
L
K
-.377
-6·0X10-13
-.543
-8·7X10-13
-.849
-13·6X10-13
-1·508
-24·1X10-13
-3·394
-54·3X10-13
-13·56
-217 X 10-13 ergs
Lyman Series
Balmer Series
Paschen Series
Brackett Series
one electron-volt, the energy acquired by an electronic charge in falling through a potential difference of one volt.
= 4·803 x 10-10 / 300 = 1·6 x 10-12 ergs
Scale, Electron-Volts
0
1
2
3
4
5
6

FIG. 272.

farther than P, and P appears only for $Z = 90$ or over. O appears for $Z = 73$ and over.

The simple Bohr theory we have given does help us to understand X-ray series, in spite of the fact that it does not really work for any atom with more than one planetary electron; for though it does not quite work, it very nearly does, especially at X-ray levels. For example, the K level contains 2 electrons in every normal neutral atom except hydrogen.

Moseley and the Atomic Number

Every element from $_{11}$Na to $_{92}$U gives, as the first line of its K series, the frequency which results from a transition from the L level (to be exact, from one of the L levels—the corresponding one for each atom). This frequency so nearly resembles the one we should get from our simple Bohr theory that Moseley (who spotted the theoretical significance of Bohr's theory as soon as it appeared) was able to do, in 1913, a series of experiments which put the periodic table in order once and for all; one of the crucial series of experiments in the history of physics; and any one who understands the simple Bohr theory for hydrogen with circular orbits can understand Moseley's work, or anyway the most important part of it. We have seen in this chapter that for any 1-electron atom of nuclear charge Ze, we have for a transition from the L to the K level

$$\frac{1}{\lambda} = \frac{3}{4}RZ^2$$

where R = Rydberg's constant, $1{\cdot}097 \times 10^5$ cm.$^{-1}$.

Moseley as a first attempt tried to see if this would work for one of the K electrons, regardless of the fact that there was not only another electron at the K level, but also lots more at all the other levels.

He measured the wave-lengths of the K_α line of the characteristic X-rays of as many elements as he could get hold of (and he got hold of all but 5 of the 42 from $_{11}$Na to $_{52}$Te, and also some higher ones) and plotted $\sqrt{\nu}$ against Z. Since $\nu = \dfrac{c}{\lambda}$, this should by the equation for $\dfrac{1}{\lambda}$ give a straight line if the equation is true.

What it did give the table below shows.

$_{19}$K	15·616$_{\cdot 865}$	$_{47}$Ag	40·246$_{\cdot 896}$
$_{20}$C	16·481$_{\cdot 865}$	$_{48}$Cd	41·142$_{\cdot 897}$
$_{21}$Sc	17·346$_{\cdot 868}$	$_{49}$In	42·045$_{\cdot 909}$
$_{22}$Ti	18·214$_{\cdot 870}$	$_{50}$Sn	42·954$_{\cdot 898}$
$_{23}$Va	19·084$_{\cdot 869}$	$_{51}$Sb	43·852$_{\cdot 904}$
$_{24}$Cr	19·953	$_{52}$Te	44·756

In this table, values of $\sqrt{\frac{\nu}{R}}$, as observed, are given against two groups of six consecutive elements in different parts of the periodic table, and the differences between successive elements are noted. It can be seen at once that though it is clearly wrong to suppose that the simple equation is absolutely right, as the differences get slightly larger as Z increases, it is quite obvious that something is increasing progressively, a unit at a time. This something can only be the charge on the nucleus, or in other words the atomic number, Z, which is also the order of the element in the atomic table.

Before Moseley's work physicists had had to rely on the chemical atomic weight (actually, as we now know, a mere average of the weights of collections of isotopes), and in several instances this had left the true order doubtful. In particular, $_{18}A^{39\cdot91}$ and $_{19}K^{39\cdot096}$, $_{27}Co^{58\cdot94}$ and $_{28}Ni^{58\cdot69}$, $_{52}Te^{127\cdot5}$, and $_{53}I^{126\cdot93}$, all had their true orders settled, in each case against the atomic weight order and in favour of the chemists' order.

The fact that by plotting $\sqrt{\nu}$ against Z one gets a very slight curve instead of a straight line is to be expected, for one would expect the remaining electronic structure of the atom to modify the effective value of Z, so that one should put $(Z - S)$ instead of Z in the equation, with some sort of progressive change in S as Z changes.

We cannot here go further into this, except to remark that what we have called S is usually referred to as the "screening constant" (approximately 1·0 for all elements).

Fig. 273 shows a typical energy-level diagram. It gives a good general idea both of how these useful diagrams are constructed and of how the various *X*-ray spectra are set up. A diagram for the highest natural element $_{92}$U would show 1 K level, 3 L levels, 5 M levels (as here), 7 N levels, 5 O levels, 3 P levels, and it would give 4 sets of *X*-ray series, K, L, M, N, where Krypton has only K, L, M.

It will be noticed that not all transitions which appear to be possible actually happen. For instance, there is no transition from L_3 to K or from any of the N levels to M_2. All transitions happen within the restrictions of certain selection rules, which appeared first as arbitrary rules—the forbidden transitions were simply not found to happen—and which later found theoretical explanations, far outside the scope of this book.

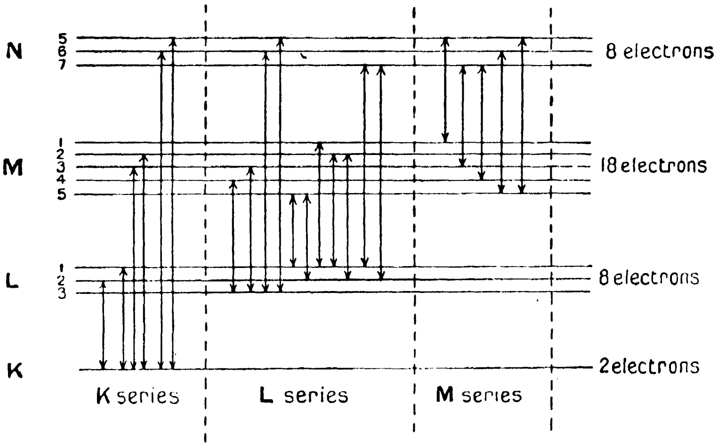


FIG. 273.

One point the table shows clearly—the explanation of the fact that above a certain threshold value of the accelerating voltage a whole series of characteristic X-rays of the bombarded metal of the target shows up, and below this none of the series. A K or L electron, or one from any level, gets kicked out of the atom only if the bombarding electron has as much energy as the negative energy of the level in question. A little less just leaves the electron where it is. However, if the electron does get out, an electron from a higher level (*i.e.* one of lower negative energy, nearer the surface of the atom) comes in to replace it, emitting an X-ray in the process. This second electron is itself replaced from a still higher level, emitting another X-ray of lower frequency, and so on. Moseley's Law works not only for the K lines but for any identifiable line whatever in any X-ray series, K, L, M, or N, which appears for more than one kind of atom.

Moseley's Law may be stated (see p. 577) as

$$\sqrt{\nu} \propto (Z - 1),$$

taking the screening constant as 1·0.

It also works for the quantum frequency corresponding to the minimum exciting voltage we have just described, since $h\nu = Ve$, and of course for characteristic absorption frequencies, as now to be described.

Absorption of X-rays

For a monochromatic beam which does not excite characteristic radiation in the absorbing material, and which is absorbed but not appreciably weakened by scattering, the discussion is simple. If a beam of initial intensity I_0 falls on a material for which the intensity I at distance x through the material differs from the intensity $(I + dI)$ at distance $(x + dx)$ through the material by an amount such that

$$\frac{dI}{I} = - a dx$$

if a is the "linear coefficient of absorption."

Integrating, and putting $I = I_0$ when $x = 0$, we get

$$I = I_0 e^{-ax}.$$

Actually it is found that for any given substance what matters is the number of atoms in unit area of the absorbing sheet, so that $\dfrac{a}{\text{density}}$ is a constant for the substance whatever its physical state. We may therefore more usefully give ax in the form μM where M is the mass of unit area of the absorbing sheet. Thus

$$I = I_0 e^{-\mu M}.$$

The amount of radiation actually absorbed depends on two factors.

First, there is the scattering. This has been found by Compton to obey a very odd law, based on the bold assumption that when a quantum hits an electron the laws of conservation of energy and momentum are both obeyed, just as if two billiard-balls were colliding, except that the collision is perfectly elastic.

It can be shown that if the quantum bounces off so that its new direction makes an angle of θ with its old direction, its

energy is reduced by an amount such that the increase in wave-length of its quantum is given by

$$\delta\lambda = \frac{2h}{mc}\sin^2\frac{\theta}{2} = 0{\cdot}0484\ \sin^2\frac{\theta}{2}\ \text{Å}.$$

or about 0·0242 Å for X-rays scattered through 90°. This does not happen to all the scattered rays. Part are scattered with unaltered wave-length and part with modified wave-length; the lighter the element the more the " Comptonized "

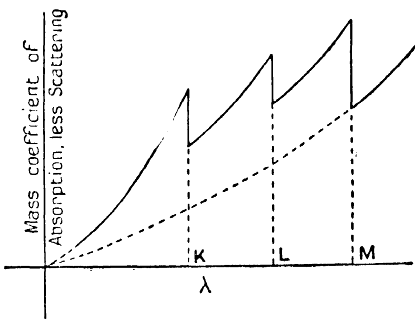


FIG. 274.

rays. (Compton's theory has been replaced by that of Klein and Nishina, based on wave-mechanics (1929).)

Secondly, there is the true absorption. This is simply the photoelectric effect, and so results in the complete disappearance of the quantum, not merely a change in its direction, as in scattering. It is found experimentally to vary with λ^3. Thus μ consists of two factors, the scattering factor which is about 0·2 for light atoms scattering medium-length X-rays, and the absorption factor which varies with λ^3.

The sort of curve one would expect for a heavy element giving K, L, and M series of X-rays would be as shown in Fig. 274.

All the curves are bits of cubics, and the values of λ corresponding to the absorption limits give quanta having energies due to the K, L, and M energy-levels in the atom. Thus K gives a single absorption limit, but the others, if closely examined, are found to be multiple, as would be expected from the fact that the L and M levels are multiple. If we write ν for the frequency of the absorption limit of a given energy

level, $\sqrt{\frac{\nu}{R}}$ increases progressively with the atomic number Z according to ` oseley's Law, just as it does for particular lines such as the K_α line as described on p. 577.

Diffraction of X-rays by Crystals

As early as 1899, attempts to produce diffraction by *X*-rays passing through a narrow slit had given a rough indication of a wave-length of the order of 10^{-8} cm., and Sommerfeld in 1912 estimated 4×10^{-9} cm. as the wave-length of the hardest *X*-rays he was using, but accurate work on *X*-ray wave-lengths did not begin till just before the 1914–18 war. In 1913 it occurred to Laue that a crystal might act as a grating, because the spacing between individual atoms in a crystal is of the same order of size as the suspected *X*-ray wave-lengths as already estimated. The work was done in Laue's laboratory with the help of his assistants Friedrich and Knipping.

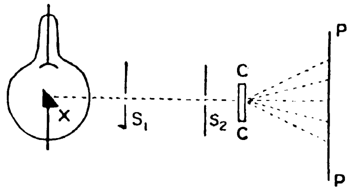


FIG. 275.

They set up an arrangement as in Fig. 275, not quite knowing what to expect. A narrow beam of *X*-rays was sent through two small holes S_1 S_2 and through a crystal (at first of zinc sulphide) to a photographic plate PP. They rather expected some sort of interference fringes to show, as they do when converging light is sent through half-silvered parallel plates of glass, but what they actually got for the zinc sulphide crystal was a pattern of spots quite widely grouped about the central spot (Fig. 276).

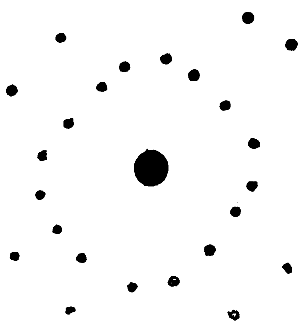


FIG. 276.

This actually was a reflected, not a transmitted, interference pattern, and was explained as such by W. L. Bragg, who got the idea when lying in his bath in a bathroom with a pattern of spots on the wall. Each of the spots was really a reflection of

the central spot as produced in one or other of the numerous planes of atoms which can be produced in a crystal as lines of dots may be produced, as in Fig. 276 A, in a collection of regularly spaced dots. It will be seen that line 1 has 16 dots per 8 cm., line 2 has 4+ dots, line 3 has 5, line 4 has 7, line 5 has nearly $4\frac{1}{2}$, line 6 has $11\frac{1}{2}$. In the same way the planes in a crystal would have different numbers of atoms per unit area, and the more the atoms the better the plane would reflect. This brilliant experiment showed us simultaneously how to measure the wave-length of X-rays and how to investigate the structure of crystals. W. L. Bragg, with his father Sir W. H. Bragg, then devised a more powerful way of using a crystal to investigate X-rays.

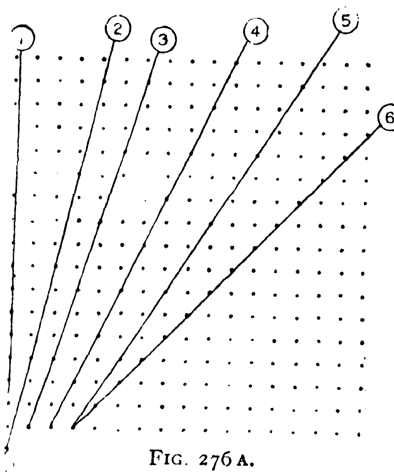


FIG. 276 A.

They set up an arrangement developed from Laue's in which the beam of X-rays was reflected from a crystal

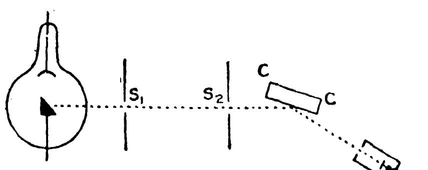


FIG. 277

and was caught by an ionization chamber instead of a photographic plate (Fig. 277). The apparatus was arranged so that the angle through which the beam was bent by the crystal could be varied, the angles of incidence and reflection being kept the same by moving the crystal. The ionization-current was then plotted against the angle of deflection, and a result of the type of Fig. 278 was obtained.

There would, of course, be a falling off in intensity as the deflection angle increased, but the same figure would recur if

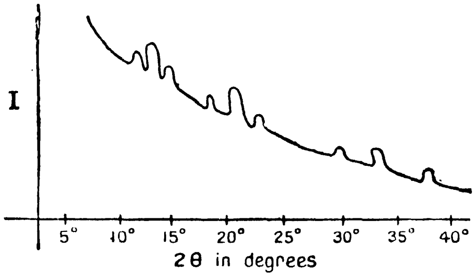


FIG. 278.

the characteristic radiation of the target was superposed on the continuous radiation from the source.

The maxima were obtained simply by the reinforcing effect when the path difference for the rays from successive reflecting planes in the crystal was a whole-number multiple of the wave-length of the X-ray (Fig. 279).

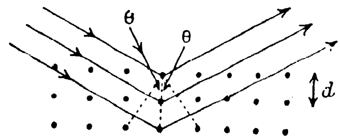


FIG. 279

Thus if
$$n\lambda = 2d \sin \theta$$
where d is the distance between the crystal planes, there will

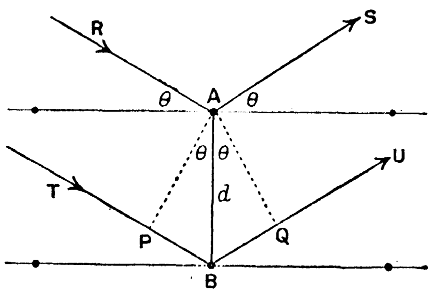


FIG. 280.

be reinforcement and the ionization current will show a peak.

For suppose RAS and TBU (Fig. 280) to be rays incident on successive crystal planes containing crystals A and B. The

rays will reinforce one another when the extra distance travelled (PBQ) is a whole number of wave-lengths.

Thus the condition for a peak is

$$n\lambda = \mathrm{PBQ}$$
$$= 2d \sin \theta$$

from the geometry of the figure.

Radioactivity and the Nucleus

Consider the nucleus of a heavy atom as a complex structure with energy-levels, but with a more complicated law of force than that which obtains in the outer electronic structure. Suppose that some kinds of nucleus are unstable so that for a large number N of one particular kind of nucleus a proportion λN is found to break down every second.[1] Assume that these nuclei are built up of protons and neutrons, and that protons and neutrons tend when possible to arrange themselves in very stable groups of 4—2 protons and 2 neutrons —so that when an unstable nucleus breaks down either one of these stable units is expelled or else an electron, and that as a result of the rearrangement of energy-levels a number of photons may follow the expelled particle, just as K, L, M, or N photons are emitted as a result of the removal of one K-level electron from the outer electronic structure of an atom. Call the 4-unit particles α-particles, the emitted electrons β-particles, and the photons γ-rays (and these γ-rays will include X-rays resulting from secondary disturbances in the outer electronic structure as a result of the nuclear explosion); and you will have a good general idea what to expect from radioactive substances.

There is a point here about which we must allow no possibility of misunderstanding; a point on which the situation has changed sharply since 1932. If a particle, such as a positron, comes from a nucleus, we cannot infer that it existed as a separate entity inside the nucleus before being detected (through the ionization it causes) by means of a photographic plate.

Before 1932, the only nuclear particles known were protons (mass 1 unit, charge " + 1 " electronic unit) and electrons (mass effectively zero, charge " − 1 " electronic unit). These were considered to be grouped inside the nucleus, so far as

[1] This λ is the transformation constant, and has nothing to do with wave-length here. Unluckily λ has acquired these two unrelated meanings.

possible, in groups of 4 protons and 2 electrons, each group forming an α-particle.

As no one saw how protons, all having the same charge and presumably repelling one another, could hold together, every one considered that both protons and electrons actually existed in the nucleus.

Since 1932 many other particles have been discovered; neutrons, positrons, several kinds of mesons, neutrinos, and deuterons. Although they may be detected as nuclear products, it would be wrong to infer that these particles necessarily exist as such inside the nucleus before being produced.

A nucleus is now regarded as made up of nucleons, a term which means a charged or uncharged particle of unit mass, and which at present includes protons and neutrons. α-particles are thus groups of 4 nucleons (2 protons and 2 neutrons), apparently a particularly stable combination, which seems to come out in every spontaneous radioactive change involving loss of mass units. If a negative proton is discovered, it will presumably also count as a nucleon. On the scale on which O^{16} is 16·00000 mass units, the neutron mass is, in 1949, taken as 1·00893, the proton 1·00758, and the electron (rest-mass) 0·00055. Nevertheless it is supposed that a proton may lose a positron and become a neutron or gain an electron and become a neutron; and that a neutron may gain a positron and become a proton or lose an electron and become a proton. The question of how a proton can apparently lose a positron and become heavier, or a neutron gain a positron and become lighter, is involved with considerations of the exchange of mass and energy which are discussed later.

When a secondary photon is emitted by the outer electronic structure of an atom as a consequence of the radioactive breakdown of the nucleus, this photon is called a γ-ray, not an X-ray, although there is no real difference in nature. There is, however, a difference in behaviour. The cathode rays causing X-rays may have a continuous range of velocities, and the resulting primary X-rays a continuous range of frequencies. The β-rays which correspond to the cathode rays may also have a continuous range of velocities, but the primary γ-rays from the nucleus do *not* have a corresponding continuous range of frequencies. They seem to result from rearrangements among the energy-levels within the nucleus. Suppose a heavy atom,

is unstable. It may readjust itself by expelling either an electron (or β-particle) which leaves its mass unchanged but raises the nuclear charge by one unit, or else an α-particle (or helium nucleus) which lowers its mass by 4 units and its charge by 2 units. This explosion may produce an atom which is also unstable, and the same process repeats itself until a stable product is reached. In each of the three known natural radioactive chains, this final product is a form of lead.

Radioactivity was discovered by Becquerel in 1896, partly as a result of Röntgen's discovery of X-rays the year before; for it had put physicists on the look-out for penetrating radiations. A salt of uranium was found to phosphoresce, and was then found to affect a photographic plate through black paper and later through a thin sheet of silver. The phosphorescence was at first thought to be the cause, and it was to investigate this idea that the experiments were carried out; but it was found that the phosphorescence was due to some other component, but the penetrating radiation came from the uranium. M. and Mme. Curie pursued the matter and discovered the far more powerfully radioactive radium, and Rutherford and Soddy began a long series of investigations.

Radioactive Series

Before going further, the reader should study the chart at the end of the book in some detail. The general idea is simple, but the chart does want careful study. Atomic numbers are plotted vertically, mass numbers horizontally. Along each line representing an α transition is given the half-life of the element at the top end, and the energy in MeV, or range in centimetres of air at N.T.P. of the α-particle giving this transition. The half-life (see p. 587) is the time taken for half the atoms in any sample of the material to suffer transformation, and is given in years, months, days, hours, minutes, or seconds, as the case may be.

There are only four groups of series known, which are separate because all mass changes are due to α-particles of mass 4 units and charge 2 units.

The series beginning with ${}_{92}U^{238}$ is known as the $(4n + 2)$ series. If you divide 4 into 238, or into any member of this series, the remainder is 2. This series is coloured blue on the chart. The main natural series is shown with double

lines. The single-line and dotted series are recently (in 1949) discovered subsidiary series arising from products produced artificially.

The thorium ($4n$) series is coloured green, and the third natural series ($4n + 3$), often known as the actinium series but derived from $_{92}U^{235}$, is coloured purple. The red ($4n + 1$) series is recently discovered, and is known as the neptunium series, its longest-lived member being the artificially produced $_{93}Np^{237}$.

The chart should explain itself. The meanings of the Geiger-Nuttall diagram, the relation between range and energy, and the disintegration constant λ are explained in this chapter, and the artificial transitions and the origin of the ($4n + 1$) series in Chapter VI. There are four other naturally occurring radioactive isotopes (1949).

A N	*Element*	*A. W.*	λ	*Half-life*	*Particle*	*Range (cm.)*
19	Potassium	40	$1{\cdot}16 \times 10^{-17}$	$1{\cdot}9 \times 10^{9}$ years	β	
37	Rubidium	87	$4{\cdot}4 \times 10^{-19}$	5×10^{10} years	β	
62	Samarium	148	$1{\cdot}57 \times 10^{-19}$	$1{\cdot}4 \times 10^{11}$ years	α	1·15
71	Lutecium	176	$3{\cdot}14 \times 10^{-19}$	7×10^{10} years	β	

It has recently (1948) been found that the air contains a small but constant proportion of C^{14} atoms, β emitters with a half-life of about 5000 years, no doubt due to neutron bombardment associated with cosmic rays. These may prove useful in archaeological research, because the proportion of C^{14} to ordinary carbon in living breathing creatures is the same as that in the air, but in dead animals which have ceased to breathe it gradually diminishes with time. This fact may enable us to date approximately, but reliably, the animal remains of the not too remote past.

Half-life and Average Life

For every radioactive element known, the proportion of its atoms breaking down per second is absolutely invariable, and

we will now follow up the mathematical consequences of this fact.

Suppose that out of every N atoms of a particular substance a proportion λN change in every second. Then

$$dN = -\lambda N dt.$$

Integrating this equation and putting $N = N_0$ when $t = 0$ we get

$$N = N_0 e^{-\lambda t}.$$

The half-value period is the time it takes for N to sink to half as much as N_0.

This is when $$\frac{N}{N_0} = \frac{1}{2} = e^{-\lambda t}.$$

So that, taking logs on each side

$$-\log_e 2 = -\lambda t,$$

for the half-life.

Since $$\log_e 2 = 0\cdot693,$$

we have $$\text{half-life} = \frac{0\cdot693}{\lambda} \text{ sec.}$$

λ then is always measured in $(\text{sec.})^{-1}$.

The half-life must not be confused with the average life. The half-life is the time it takes any collection of atoms to be reduced to half that number. The average life is what you get by adding together the whole lives of all the atoms concerned and dividing by the total number of the atoms concerned.

Thus for the average life $\bar{t}$, we have

$$\bar{t} = \frac{1}{N_0}\int_{t=0}^{t=\infty} t\,dN$$

$$= \frac{1}{N_0}\int_0^{\infty} -tN_0\lambda e^{-\lambda t}dt$$

(since $$N = N_0 e^{-\lambda t} \text{ and } dN = -\lambda N_0 e^{-\lambda t}).$$

$$\therefore \quad \bar{t} = -\lambda\int_0^{\infty} te^{-\lambda t}dt$$

from which, integrating by parts,

$$\bar{t} = -\lambda\left[-\frac{1}{\lambda^2}\right] = \frac{1}{\lambda}.$$

Thus the average life is always $\frac{1}{0{\cdot}693}$ or 1·443 of the half-life.

The problem is not often so simple. In general, a whole string of radioactive substances is involved in any activity; and here and now we will only consider cases of equilibrium, where the members of a chain have settled down to a condition in which each is being produced just as fast as it is decaying. In these circumstances it is obvious that $\lambda_1 N_1 = \lambda_2 N_2 = \lambda_3 N_3 = \ldots$; and, since the half-lives are inversely as λ, it follows that the amount of each product present (except the stable one at the end, which just goes on growing) is proportional to the half-life.

This fact gives us the best way to determine both very long half-lives (as of Uranium I) and the age of the earth. Consider, for example, the uranium chain. Suppose U^{238} is found in equilibrium with its products in a rock. The amount of each product is measured. The half-lives of some of the shorter-lived ones are found by direct experiment, and compared as a cross-check. The relative abundance of these and the medium-lived ones is found, and the half-lives of these calculated. From these we go to the long-lived ones. The amount of lead present per unit of Uranium I gives the time that the uranium has been embedded in the rock. The age of the lower carboniferous has been estimated as 340,000,000 years by this method.

Radioactive Radiations, α, β, γ

The three natural types of radioactive radiation show wide differences in range, energy, velocity, specific ionization, and total ionization.

In range, for rays of similar energy, where α-particles are stopped by a few centimetres of air at N.T.P., or by a sheet of paper, β-particles will penetrate, say, 20 metres of air or a few millimetres of aluminium, and γ-rays several inches of lead. (Of course, γ-rays do their penetration in a different way. α-particles and β-particles, being particles, either arrive or do not; there is a definite point where the particle is finally stopped).

γ-rays, being quanta of electromagnetic waves (or photons), disappear catastrophically when they disappear at all, and hence a group of γ-ray quanta falls off exponentially with the thickness of material penetrated. α-rays lose their energy

gradually by collisions along their path, and hence cover a distance which nearly always falls between rather narrow limits before being stopped.

There is, of course, no real difference between γ-rays and X-rays. The custom has grown up of giving the name X-ray to short electromagnetic waves from the outer electronic structure of the atoms, and of γ-rays to precisely similar waves from the nucleus. Thus what we have said about absorption of X-rays holds for γ-rays. In velocity, α-particles range from about $\frac{1}{15}$ to $\frac{1}{25}$ of the velocity of light; β-particles reach up to 0·999 (and more) of the velocity of light; and γ-rays, of course, are a kind of light, and so travel with the velocity of light.

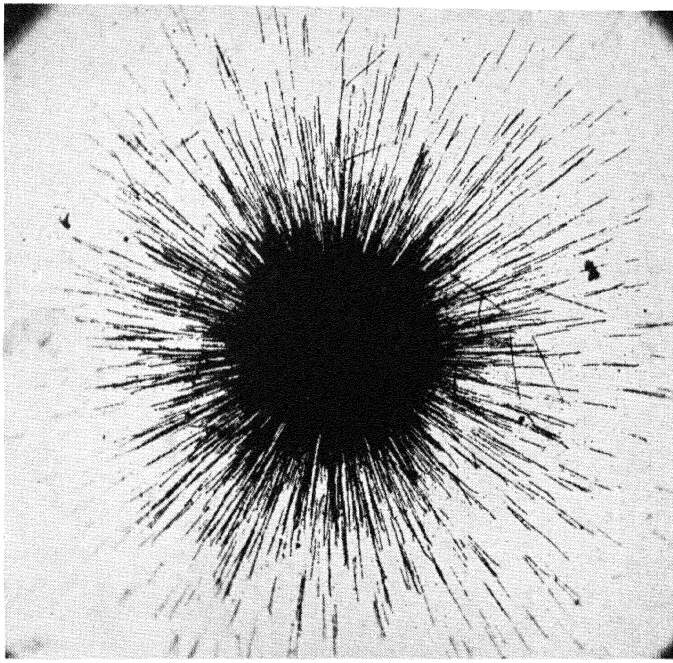
The specific ionization (number of ion pairs formed per centimetre of path) is much greater for α-particles than for β-particles. A 3-MeV α-particle produces about 4000 ion pairs per millimetre in air at N.T.P., and a β-particle of the same energy only about 4.

Given the same initial energy, an α-particle, a β-particle, and a γ-ray photon produce about the same total ionization. With an α-particle this is almost entirely concentrated in a very much shorter path, as not only is its path relatively very short but also the specific ionization increases near the end very rapidly as the particle slows down. The ionization by γ-ray photons is entirely through the secondary electrons they produce, either as photoelectrons or as Compton electrons.

α-Rays.—α-particles have the convenient power of causing scintillations on a fluorescent screen. You can see them with a microscope by looking at the dial of a luminous watch. The number emitted by a known weight of a given radioactive material can be determined by counting the scintillations on a screen at a convenient distance subtending a known solid angle. Their charge is found by allowing a known number to strike a conductor attached to an electrometer, and is found to be two electron units each. Their nature was originally found by allowing them to penetrate a thin glass tube into an exhausted chamber, which was then found to contain helium. Their $\frac{e}{m}$ ratio was found by bending them in magnetic fields—and extensive fields of considerable strength are needed. Peter Kapitza was the first to produce them in the Cavendish laboratory in the 1920s. The velocity is found by these strong fields (as it is for electrons with weaker fields) and

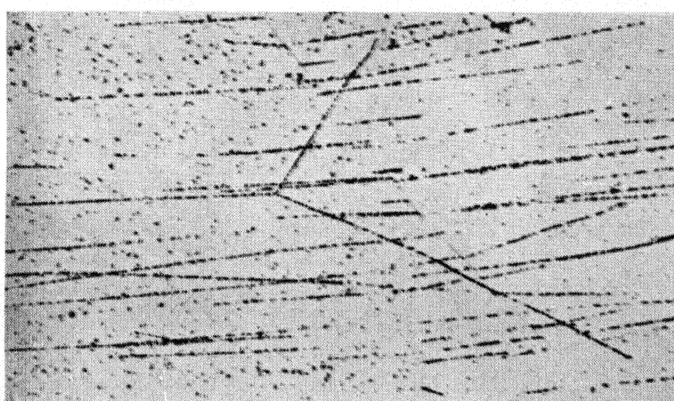
Plate 3

Below: Tracks of α-particles from a speck of radium



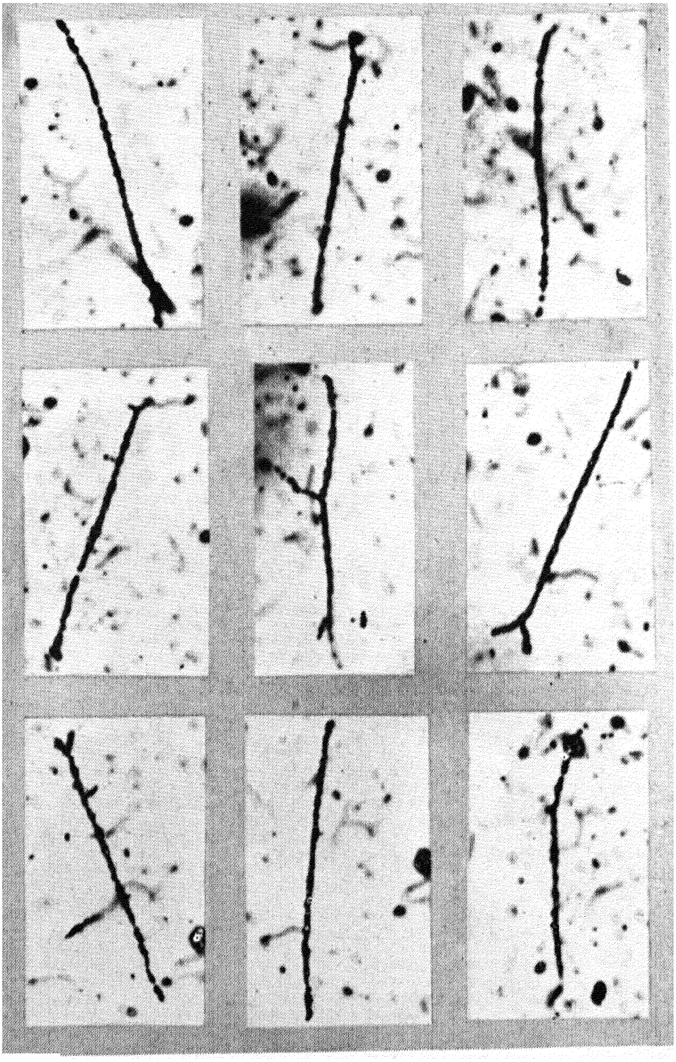
← - - - - 100 μ - - - - →

Below: Proton-deuteron collision



← - - - - - - - - - 100μ - - - - - - - - - →

Plate 4



← - - - - - - - 25 μ - - - - - - - →

Fission tracks due to fission of U^{235} by slow neutrons

m is found from e and $\frac{m}{e}$. It is always the mass of a helium (this method is not accurate enough to distinguish between the mass of a helium nucleus and that of a helium atom).

The Geiger-Nuttall Relation

Interesting and fairly exact mathematical relations are found between the range, velocity, and ionizing power of α-particles.

Ionizing power is believed to be proportional to the electrostatic attraction between the α-particle and the electron, and to the time for which it is effective. The former is constant, the latter $\propto \frac{1}{v}$, where v is the velocity. Hence it is not surprising to find that the ionizing power increases enormously toward the end of an α-particle's range.

The velocity of the α-particle and the amount of range R that it still has to cover are related by $v^3 = \mathrm{KR}$, where K is a constant. The decay constant λ (which varies inversely as the half-life) is connected with the initial range R_0 by Geiger and Nuttall's equation

$$\log \lambda = \mathrm{A} + \mathrm{B} \log \mathrm{R}_0$$

where A and B are constants. It is found that B is the same for the three natural radioactive chains, but A is slightly different for each.

Since if $v^3 = \mathrm{KR}$, we have that $3 \log v = \log \mathrm{K} + \log \mathrm{R}$, we can also say that

$$\log \lambda = \mathrm{C} + \mathrm{D} \log v_0$$

where C and D are corresponding constants.

α-Ray Spectra

Some α-emitters have a spectrum of velocities for their α-particles — radioactinium has eleven components and thorium C has five—but some only have one. The balance of energy is successfully traced to γ-rays, and this relation between α-particles and γ-rays tells us something useful about the structure of the nuclei concerned. The two shortest-lived elements, Th.C' and Ra.C', emit a few particles of very long range compared with their normal ranges of 8·62 cm. and

6·95 cm. respectively (themselves longer than those of any other substance). Th.C′ emits about 224 long-range particles in every million, 34 with a range of about 9·7 cm. and 190 with a range of about 11·6 cm. Ra.C′ emits about 30 long-range particles per million, in twelve different groups with the biggest range about 11·5 cm. The fact that Th.C′ and Ra.C′ are the only substances to produce these long-range α-particles is related to the very short lives of their ordinary α-particles. The corresponding Ac.C′ has a much longer life, though still very short.

β-Ray Spectra

The methods which find e, $\frac{e}{m}$, m, and v for cathode rays can be adapted for β-rays, since they are electrons just the same, except that their enormously higher speeds—up to 0·99 of the velocity of light, and over—require corrections for the relativity change of mass with speed.

The relativity expression for the kinetic energy E_k of a particle moving with velocity v is

$$E_k = m_0c^2 \left(\frac{1}{\sqrt{1 - \beta^2}} - 1 \right)$$

if $\beta = \frac{v}{c}$ and m_0 is the rest mass.

If $\beta = 0·99$, $\beta^2 \doteq 0·98$, $(1-\beta^2) \doteq \frac{1}{50}$, and $\frac{1}{\sqrt{1 - \beta^2}} \doteq 7$.

So we have

$$\begin{aligned} E_k &\doteq 6m_0c^2 \\ &\doteq 6m_0v^2 \text{ [since } v = 0·99\,c] \\ &\doteq 12 \times (\tfrac{1}{2}m_0v^2). \end{aligned}$$

So that the K.E. is about twelve times as great as it would come on the ordinary $\frac{1}{2}mv^2$ formula.

In ergs, this comes to $6 \times 9 \times 10^{-28} \times (3 \times 10^{10})^2$, very nearly, taking approximate values of m_0 and v ($= 0·99\,c$), and this gives about 5×10^{-6} ergs.

One electron-volt is $1·6 \times 10^{-12}$ ergs, and one million-electron-volt, or MeV, is therefore $1·6 \times 10^{-6}$ ergs, so that a 3-MeV electron, a not very rare occurrence in β-rays, has a velocity of about 0·99 of the velocity of light. But an electron's energy approaches infinity as its speed approaches

the maximum possible, which is, of course, the speed of light. So at such speeds a relatively minute increase of speed produces an enormous increase in energy. For an electron of 11 MeV v comes out at about 0·9990 c.

If $v = \frac{1}{2}c$, as it does for the slower β-particles, we have

$$E_k = m_0c^2 \left(\frac{2}{\sqrt{3}} - 1\right) = 0{\cdot}1545\ m_0c^2.$$

This is $\frac{0{\cdot}1545}{0{\cdot}125}$, or 1·236 as great as the value of $\frac{1}{2}m_0v^2$, since

$$\frac{1}{2}m_0v^2 = \frac{1}{2}m_0\left(\frac{c}{2}\right)^2 = \frac{1}{8}m_0c^2.$$

It is the energy of an electron of only about 0·08 MeV. At lower values of v we can do the calculation best by expanding by the Binomial Theorem, for

$$\frac{1}{\sqrt{1-\beta^2}} = (1-\beta^2)^{-\frac{1}{2}} = 1 + (-\tfrac{1}{2})(-\beta^2) + \frac{(-\frac{1}{2})(-\frac{3}{2})}{\lfloor 2}(-\beta^2)^2 + \ldots$$

$$= 1 + \tfrac{1}{2}\beta^2 + \tfrac{3}{8}\beta^4 + \ldots$$

When $v = \frac{1}{20}c$, as for α-particles, $\beta = \frac{1}{20}$, $\beta^2 = \frac{1}{400}$, and β^4 is negligible. We may therefore write

$$E_k = m_0c^2\left(1 + \frac{1}{2}\beta^2 - 1\right) = \frac{1}{2}\beta^2 m_0c^2$$

$$= \frac{1}{2}m_0\frac{v^2}{c^2} \times c^2 = \frac{1}{2}m_0v^2.$$

This is the ordinary formula for kinetic energy, so we see how this formula emerges from the relativity formula even at speeds of 10,000 miles per second. Actually a fast α-particle, with a speed of 12,000 miles per second, has its energy increased above the $\frac{1}{2}m_0v^2$ level by only $\frac{1}{16{,}500}$ of that value. The mass of α-particles is never effectively altered by any relativity correction.

The β-ray continuous-spectrum rays are known to come from the nucleus. The number of β-particles emitted by a piece of radioactive material in a given time can be found by catching them in an electroscope and measuring the charge they carry, and the number of atoms in the source disintegrating in the same time can be calculated from λ and the mass of the piece of material. These two numbers are found

to correspond, and this makes it vanishingly improbable that the electrons come from anywhere but the nucleus.

It is equally certain that the line-spectrum β-particles come from the outer electronic structure, knocked out by γ-ray photons from the nucleus. Their energies fit this theory nicely.

The energies of the continuous-spectrum β-particles, on the other hand, do not fit the simple theory at all. One would expect that particles from the nucleus would be emitted with an energy corresponding to some definite nuclear reaction or reactions. Finding that the distribution of energies is in fact continuous (for Ra.E the high-energy limit is 1·2 MeV and the average energy 0·34 MeV) one would next suppose that the electrons with energies below the maximum had lost some in getting out. Experiment shows this is not so. All the heat of the process, including that (if any) robbed from the electron in getting out, would be caught by the calorimeter. None is so caught. If the principle of the conservation of energy holds, it would appear that some energy escapes with a very penetrating particle of no charge and sub-electronic mass, which cannot give up energy to the calorimeter. This supposed particle is named the neutrino, and indirect evidence that it exists has come from other experimental work, but there is no direct evidence, and at present, 1949, it is difficult to see how there could be.

γ-Ray Spectra

All the γ-rays emitted have their wave-lengths in definite line-spectra. Some come from the nucleus—the harder ones —and there is convincing evidence that they are not the direct consequence of the emerging β-ray, but are the slightly later result of the energy-level rearrangements that follow the emergence of the β-ray. The β-ray usually leaves the nucleus in an intermediate energy-state, and it then falls to a lower state, emitting γ-rays. Sometimes no γ-ray follows the β-ray, and it is then thought that the β-ray left the nucleus in the lowest available and therefore stable energy-state. Other, presumably secondary, γ-rays unmistakably come from the outer electronic structure, for their wave-lengths are the same as the wave-lengths of the characteristic X-rays for the element. Presumably the sequence here is that the nuclear disturbance ejects an electron from the

outer electronic structure, and the secondary γ-rays are really X-rays resulting from the re-settlement after this ionization.

γ-rays can be made (with some difficulty owing to their relatively feeble intensity) to show all the effects of X-rays (*e.g.* interference and diffraction), and their wave-length can be measured by crystals, but to do this is difficult, and it is usually inferred from the energy of the photo-electrons they can be made to kick out of appropriate atoms. Then $ev = h\nu - E_k$, where E_k is the binding energy of the K-shell, assuming that the electron comes from the K-shell, and the wave-length $\lambda = \frac{c}{\nu}$ can be found from the measured velocity, since $\lambda = \frac{hc}{ev}$.

It can be readily seen how the whole mass of the evidence requires a nucleus with energy-levels to explain it.

It used to be supposed that all β-rays and no α-rays were accompanied by γ-rays, but it is now known that some β-rays are not accompanied by γ-rays.

There is evidence that all γ-rays emitted by naturally radioactive elements are in association with α-rays or β-rays which come not from the atom emitting the α-rays or β-rays, but from its successor atom. The γ-ray seems to come about 10^{-12} or 10^{-13} seconds later, apparently because the newly formed nucleus begins in an unstable state and has to suffer an internal rearrangement to get stable. Some artificially formed nuclei do apparently emit γ-rays without α-rays or β-rays (e.g. gold illuminated by X-rays of photon energy 1·2 MeV, which emits γ-rays with a half-life of 7·5 seconds). The softest, or longest wave-length, γ-rays have a wave-length of nearly 4 Å, corresponding to a photon energy of about 0·03 MeV, and the hardest from natural sources have a wave-length of less than 0·05 Å, corresponding to an energy of 2·62 MeV. Some X-ray tubes have recently (1948) been designed to give a wave-length much shorter than that of the shortest γ-rays.

The Geiger-Müller Counter

The Geiger-Müller counter (Geiger, 1913, and Geiger and Müller, 1928) is an immensely useful and widely used piece of apparatus for detecting the presence of radioactive substances

and measuring their radiations. Its general scheme is shown in Fig. 281, more or less diagrammatically.

It is simply a low-pressure gas-container with metal walls CC which form a cathode. The anode is a wire A entering through an insulating cap. A tube T which can be sealed off can be used to change the pressure or the contained gas. The potential difference between C and A is maintained by a

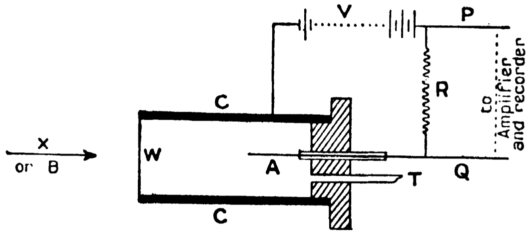


FIG. 281

potential source V, through a high resistance R. W is a gas-tight window whose thickness depends upon the nature of the radiation to be measured. Various gases are used at various pressures, but 1 cm. Hg of ethyl alcohol vapour to 10 cm. of argon is a common mixture.

The general idea is that when an ionizing particle penetrates W a current flows between A and C. This current builds up a potential difference across R with two results. (*a*) This rise can be recorded by a suitable instrument, *e.g.* a click, a lamp, or a mechanical register connected between P and Q, terminals at the two ends of R_N, and (*b*) the effective potential difference between C and A is temporarily reduced because the E.M.F. across R opposes that of V, and the current is quenched almost at once, so that the full voltage V is built up again ready for the next particle.

Fig. 281 illustrates a common design of G.-M. counter, having a metal cylinder with an insulating cap, probably of glass. It is also possible to have a metal cylinder with glass ends, or to enclose the whole of both electrodes in a gas-tight glass container, and so avoid the difficulty of a metal-to-glass seal. The size is usually about that of a radio valve—say 10 cm. long—but for certain purposes cylinders up to about 100 cm. long are used.

Fig. 282 shows the general shape of the characteristic curve, which has four well-defined sections, OA, AB, BC, CD.

Ionizing particles entering the window are called primary particles.

A few anode volts merely cause these particles to be collected, giving a negligible current. At perhaps 50 volts extra, or secondary, ions begin to be produced by collision, and each

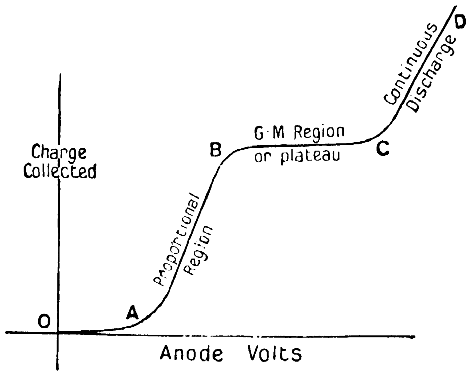


FIG. 282

collision produces one pair of ions. These ions are in turn accelerated, and at a sufficiently high voltage "avalanches" begin to occur, and we move on to AB, in which the anode volts are of the order of 500 and the primary charges are amplified about, say, 10^6 times, but over a fairly wide range of voltages—say 500 to 1000—the charge collected remains proportional to the number of ions directly formed by the entering particle. This is quite useful, because it helps one to identify the kind of primary ion producing its effect, but the apparatus is not yet a G.-M. counter. It becomes one when the voltage is high enough to get us to the region BC, called the Geiger-Müller region or the plateau (beginning at about 1000 volts), but not high enough to produce a continuous discharge (CD on the curve).

Every counter has a background rate of pulse-production when no specific radioactive sources are present. This is due partly to local contaminations and radioactive sources in the ground, and partly to cosmic radiation. Local sources can be almost entirely cut out by screens of non-radioactive lead

a few centimetres thick which do not appreciably reduce the effect of the exceedingly penetrating cosmic rays. The total background rate is from 1 to 1½ pulses per minute per sq. cm. of the longitudinal cross-section of the sensitive part of the counter. The shielding shows that the local sources contribute about half of this, and the cosmic rays the other half.

Lastly, I cannot refrain from quoting Pollard and Davidson's *Applied Nuclear Physics* about counters.

"At the Massachusetts Institute of Technology the visitor is shown a Geiger counter consisting of a fork and spoon in an evacuated space. It works! On the other hand, the reader may well intentionally make up a Geiger counter after the best instructions and fail to make it work. The authors would like to meet the counter expert who has not at some time in his career found he has constructed a 'lemon.'

"There is . . . a considerable amount of trial and error about counter construction. Of three similar counters, one may completely fail to operate for no clear reason. If, however, one thinks of the manner of operation of the counter, this behaviour is to be expected to some extent, for the counter must be such that it will discharge at a certain voltage when 'tripped' by an entering electron, and yet extinguish when the applied voltage falls by 100 volts or so. Since the nature of the discharge in the counter depends greatly on the surface of the electrodes, any dirt may or may not change the behaviour of the counter in a radical way. Moreover, the gas used for filling plays an important role and must be chosen with care. There is, therefore, in the authors' opinion, no infallible set of rules which when followed will guarantee perfect operation of a counter; the proof of the pudding is in the eating. Preparation of Geiger counters is about on a par with cooking a meal; experience makes a difference. Once a counter is found which operates satisfactorily and does not contain reactive vapours and has no leak, it will continue to operate indefinitely. In several laboratories as many as fifty counters are in operation at one time, giving no trouble, so that they are not uncontrollably temperamental."

As a footnote to this, and just to show how the fun may go out of physics, I may add that now (1949) G.-M. counters are produced and sold with a guaranteed sensitivity and life.

CHAPTER V

PHOTO-ELECTRICITY, THERMIONICS, VALVES, ACCELERATORS

Photo-electricity and Einstein's Equation—Thermionics and Richardson's Equation—Electrons in Metals—Photo-electric Cells—Valves—The Diode—The Triode—Voltage Amplification by the Triode—Accelerators, the Cyclotron, Betatron, Synchrotron, and Van de Graaff Generator.

Photo-electricity and Einstein's Equation

HERTZ, the discoverer of wireless waves (which he was searching for), discovered the photo-electric effect by accident. Both discoveries were made in 1887, and in the same piece of work. He noticed that the spark between the terminals of his detecting circuit passed more easily when these terminals were illuminated by the primary spark than when an opaque screen was placed between the two, and also when a sheet of glass was there instead of the screen. Hertz correctly inferred that ultra-violet light caused the phenomenon, because it was the only agency which could be selectively stopped by the glass. Hallwachs, three years later, showed that a negatively charged plate could be discharged in this way, but not a positively charged one, and during the next two years it was shown that in a circuit as shown in Fig. 283, in which ultra-violet light falls both on the wire gauze and on the plate, current can be made to flow so that electrons flow from the plate to the gauze, but no appreciable current can be made to flow in the opposite direction.

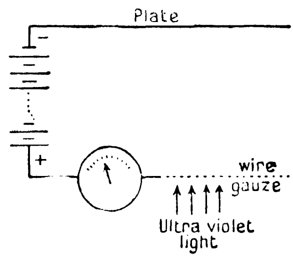


FIG. 283.

The effect was not explained till the electron was discovered in 1897, and this discovery made it probable that electrons were being jerked out of the material in some way by the ultra-violet light. This was definitely proved by Lenard in 1900.

He found $\frac{e}{m}$ for the emitted particles by bending them in a magnetic field, and by applying small opposing electric fields proved that some of them at least did not merely leak out of the material, but were projected with an energy of about two electron-volts.

A number of other facts then emerged about photo-electrons. The size of the saturation current (the current with enough applied electric field to collect all the available photo-electrons coming out of the plate) is proportional to the intensity of the light that causes it, but the energy of the fastest electrons emitted has nothing whatever to do with this intensity, but depends solely on the wave-length of the shortest waves of the illuminating light. There is a threshold minimum wave-length required, for any given pure material, to bring any photo-electrons out at all. For most elements, a large change in the temperature of the emitter has little effect. Electrons begin to come off almost instantaneously when any illumination, of sufficiently short wave-length, no matter how feeble, is switched on.

It was this last fact which gave the most easily intelligible blow to the classical theory of radiation, according to which energy passes continuously. According to classical theory it was calculated that for some elements a light, of so feeble an intensity that one atom would take over a year to accumulate the energy it would need for a photo-electron, caused instantaneous emission of photo-electrons of maximum energy. Clearly some drastic new theory was needed, and it was supplied by Einstein, using Planck's Quantum Theory, in 1905.

His theory completely accounted for all the phenomena then known. He supposed that the radiation arrived in quanta of energy $h\nu$, and that if any energy was given up a whole quantum could be given up, and this could happen instantaneously, as soon as the light was switched on. The energy delivered would depend on the frequency (or wave-length since $\lambda\nu = c$) and on nothing else, so that it would be absolutely independent of the intensity.

Though the word photon, for a quantum of radiation travelling through space, is much more recent than Einstein's development of the Quantum Theory, it embodies the idea that Einstein already had, of radiant energy being made up of

particles of a kind having not only energy but also momentum. One of the ways in which the Einstein quantum, or photon, was distinguished from other particles was in its ability to give up all its energy catastrophically to a material particle, itself ceasing to exist independently.

Einstein then supposed that a certain amount of work, W, is needed to get the electron out of the surface, so that we have, for an electron of velocity v and mass m,

$$\frac{1}{2}mv^2 = h\nu - \mathrm{W}$$

since $\frac{1}{2}mv^2$ is the kinetic energy passed on to the electron.

If ν_0 is the threshold frequency we may easily rewrite this

$$\frac{1}{2}mv^2 = h\nu - h\nu_0.$$

So far, so good; but this highly successful equation produced the most thrilling, good-tempered controversy ever known in science. It was good-tempered, not only because physicists are, from the nature of their subject, the best-tempered kind of scientists, but also because the battleground was in every physicist's own heart. Experimentally it could be shown that both interference and the photo-electric effect could be produced instantaneously with light of such feeble intensity that it could contain only one quantum per cubic metre. Interference requires about two independent quanta per cubic metre. The photo-electric effect requires that the whole quantum should be captured by one atom, which means that it must be concentrated in 10^{-30} cubic metre or less. It seemed that the quantum must be equally capable of occupying either of two different volumes, one of them at leat 10^{30} as great as the other. I remember a remark of Sir William Bragg's circulating in the Cavendish laboratory in 1920. He said that for the present we must accept the wave theory on Monday, Wednesday, and Friday, and the quantum theory on Tuesday, Thursday, and Saturday.

Sunday was left out. I like to think that our present relativity - quantum - wave - mechanics came in on Sunday. Anyway, it was clear to all in 1920 that in some sense both the wave theory and the quantum (or corpuscular) theory of light must express a good deal of truth.

Thermionics and Richardson's Equation

In 1880 Elster and Geitel showed that charged bodies at red heat lose a negative charge more easily than a positive, and in 1899 J. J. Thomson measured $\frac{e}{m}$ for the emitted charged bodies, by the method he used for the cathode rays, and found the thermions, as they were called, to be also electrons. These electrons were found to be emitted with rather small velocities, and without an accelerating potential to be easily stopped by an accumulation of space-charge near the emitting surface.

With enough accelerating field near the surface to enable all the electrons to get away, there is at any given temperature a maximum rate of supply of electrons per square centimetre per second for any particular material. It increases rapidly with the absolute temperature. It is possible to deduce from pure thermodynamic reasoning what this current should be. O. W. Richardson gives the equation

$$I = AT^2\epsilon^{-\frac{e\phi}{kT}}$$

where I is the current, A a constant, T the absolute temperature, ϵ the natural base of logarithms to distinguish it from e the electronic charge, ϕ a potential-difference constant so that $e\phi$ is a quantity of energy which can be read off in electron-volts if ϕ is in volts, and k is Boltzmann's constant.

It can be shown that if R is the gas constant of the Gas Law $PV = RT$, E the energy per molecule at temperature T, and N Avogadro's number, the number of molecules in a gramme-molecule, then

$$\text{Energy per gramme-molecule} = EN = \frac{3}{2}RT.$$

Then
$$E = \frac{3R}{2N}\cdot T.$$

This $\frac{R}{N} = k$, Boltzmann's constant, and since $R = 8{\cdot}315 \times 10^7$ ergs per gramme-molecule per degree, and $N = 6{\cdot}02 \times 10^{23}$ molecules per gramme-molecule, $k = \frac{8{\cdot}315 \times 10^7}{6{\cdot}02 \times 10^{23}} = 1{\cdot}381 \times 10^{-16}$ ergs per molecule per degree.

The work-function, ϕ, varies with different materials. The constant A ought theoretically to be the same for all materials (about 60 amps per sq. cm.), but actually it seems to vary a good deal.

If logarithms of both sides of Richardson's Equation are taken, we can rewrite it

$$\log \frac{I}{T^2} = \log A - \frac{e\phi}{kT}.$$

This equation gives us a straight line if we plot $\log \frac{I}{T^2}$ against $\frac{I}{T}$, and the slope of this straight line gives us $\frac{e\phi}{k}$, so that knowing k we can find ϕ.

The point of all this is that it links up the thermionic effect and the photo-electric effect, for we can compare $h\nu_0$, the energy required to detach a photo-electron from the surface, with $e\phi$, a quantity of energy which Richardson's Theory requires to represent the energy needed to get a thermion out at absolute zero.

The results of the comparison are startling. We could either compare $h\nu_0$ in ergs with $e\phi$ in ergs, or either in electron-volts. Electron-volts tell us more than ergs by giving us a better picture of the obstacle presented by the hold of the surface on the electron, so in the table shown the figures stand for electron-volts.

Metal	$h\nu_0$	$e\phi$
Platinum	6·30	6·27
Rhodium	4·57	4·58
Tantalum	4·05	4·07
Tungsten	4·58	4·52
Gold	4·82	4·42

These figures seem to show conclusively that photo-electrons and thermions have the same difficulty in getting out of the surface, and hence come from the same place originally.

Electrons in Metals

What is this place? The answer to this question (so far as it is known at present) is so fascinating that even though it is really outside the scope of this book we will try to give it, using (as in all this section) a shortened form of the treatment given by Richtmyer and Kennard in their *Introduction to Modern Physics.*

In the classical explanation of thermal and electrical conduction in metals (due to Drude and Lorentz), the " free " electrons in a metal are treated as are molecules in a gas, with collisions, mean free paths, a Maxwellian distribution of

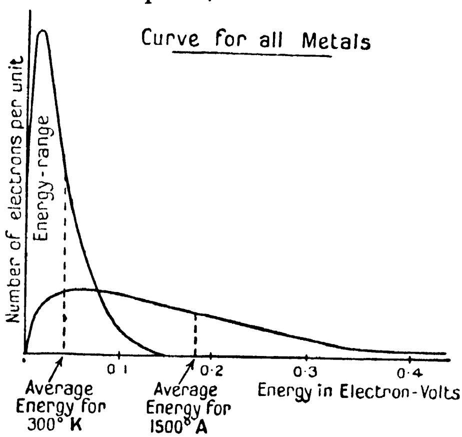


FIG. 284

velocities, and other orthodox comforts for the theorists. According to this theory there would always be some electrons in the " gas " moving fast enough to escape from the surface, and as the temperature rises the number of these should increase.

At absolute zero all the electrons should have zero kinetic energy. At any higher temperature the energies should be distributed among the electrons in a Maxwellian way, and the average energy at any absolute temperature T should be proportional to T.

The curves in Fig. 284 show how the energy is distributed among the electrons (or how the electrons are distributed

among the possible energies). There is no curve for absolute zero, because on this theory all electrons should have no energy at absolute zero.

It will be noticed that the average energy is about 0·18 ev (electron-volts) at 1500° A., and 0·036 ev at 300° A.

Compare this classical distribution with the Fermi-Dirac distribution, based on the later Quantum Theory, in Fig. 285.

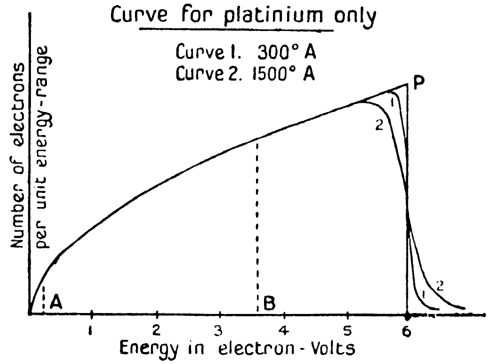


FIG. 285.

At absolute zero the curve reaches a sharp peak at P. The mean energy is then, for platinum, 3·6 ev ($\frac{3}{5}$ of the maximum energy of 5·975 ev).

The difference is so great as to be almost incredible (as often happens with Quantum-Theory interpretations of problems at which classical theory has failed). Note that the diagrams are on a quite different scale. The whole horizontal range of the classical diagram takes one less than half-way to the first division, marked 1, of the quantum diagram. The point A represents the same thing on both diagrams, an average energy of 0·18 electron-volts, the classical value for 1500° A. It corresponds for the quantum curve to the point B, representing an energy about twenty times as great, 3·6 ev! And this energy is the average (or near enough) for all temperatures from zero to 1500° A.

The energy curve for electrons in metals at 0° A. follows the main curve up to the peak P, and then falls vertically. That is to say, electrons in platinum at absolute zero have energies

up to nearly 6 ev (5·975 ev), but none with energies beyond this. Curve 1 is the curve for 300° A. The electrons now have some energy of thermal agitation added to them, and so the top corner of the curve is rounded off. Obviously the few electrons with energies above 6 ev will be robbed from those that had nearly 6 ev before. Curve 2, for 1500° A., shows the same effect, only more so. The average energies for 300° A. and 1500° A. are so little above those for 0° A. that the difference is not shown. The value of nearly 6 ev is obtained from the calculation

$$\text{Maximum energy at } 0^\circ \text{ C.} = \left(\frac{h^2}{2m}\right)\left(\frac{3n}{8\pi}\right)^{\frac{2}{3}} \text{ ergs}$$

in which h is Planck's constant, m the electron's mass, and n the number of electrons per c.c. in platinum, estimated as the same as the number of atoms ($6{\cdot}6 \times 10^{22}$).

To get the value in electron-volts one divides by $1{\cdot}601 \times 10^{-12}$ ergs, the energy of an electron-volt. This extraordinary result, worked out by Sommerfeld in 1928, solved another problem on the way. The classical electron theory, which accounted fairly well for thermal and electrical conduction, required the electron-cloud to contribute considerably to the specific heat, since the average energy of the electrons goes up in proportion to the absolute temperature. Experiment does not show any such effect. Sommerfeld's theory gives the electrons almost the same energy at all temperatures. The increase with rise of temperature is very much smaller, and no big effect on the specific heat is required.

The maximum absolute-zero energy of " free " electrons in platinum is seen to be 5·975 ev. From the table on p. 603 we see that the work-function ϕ and the threshold-energy $h\nu_0$ for platinum are about 6·30 ev. The work required to get an electron of initially zero energy out of platinum should therefore be about 12·3 ev. Let us try to get an idea of just what happens about the photo-electrons from platinum, from the energy point of view, if a quantum of ultra-violet light of frequency 800 Å, 8×10^{-6} cm., falls upon it.

This photon has energy $\dfrac{h\nu}{1{\cdot}6 \times 10^{-12}}$ electron-volts, and

$$\nu = \frac{3 \times 10^{10}}{8 \times 10^{-6}} = 3{\cdot}75 \times 10^{15},$$

so that

$$\frac{h\nu}{1{\cdot}6 \times 10^{-12}} = \frac{6{\cdot}6 \times 10^{-27} \times 3{\cdot}75 \times 10^{15}}{1{\cdot}6 \times 10^{-12}}.$$

$$= 15{\cdot}0 \text{ ev.}$$

Suppose the temperature is 27° C., 300° A. Then the average electron will have already, before absorbing the quantum, $\frac{3R}{2N}T$, or about 0·04 ev from thermal energy available. The average electron will also have 3·6 ev from Sommerfeld energy. Thus the total energy the electron will get will be 15·0 + 3·6 + 0·04 = 18·64 ev. This will have to overcome the whole work of getting out, of 6·30 + 5·97, or 12·27 ev. So the average electron should get out with

$$\begin{aligned}(18{\cdot}64 - 12{\cdot}27) &= 6{\cdot}37 \text{ ev.}\\ &= 6{\cdot}37 \times 1{\cdot}6 \text{ ergs.}\\ &= 10{\cdot}2 \times 10^{-12} \text{ ergs.}\end{aligned}$$

Its speed will then be given by

$$\frac{1}{2}mv^2 = 10{\cdot}2 \times 10^{-12}$$

$$v^2 = \frac{2 \times 10{\cdot}2 \times 10^{-12}}{9{\cdot}1 \times 10^{-28}} = 2{\cdot}24 \times 10^{16}$$

$$v \sim 1{\cdot}5 \times 10^8 \text{ cm./sec.}$$

The fastest electrons (except for a very few) will get the full 15·0 + 5·97 ev + 0·04 ev, and hence will get out with (15·0 + 6·01 − 6·3), or 14·71 ev.

The slowest will get negligibly little of either the Sommerfeld energy or the Maxwell energy and will emerge with (15·0 − 12·3), or about 2·7 ev.

The photo-electric threshold will be for a wave-length which will give enough energy to supply the remainder of what is needed by an electron which already has the *maximum* Sommerfeld energy + the *average* Maxwellian energy. This will be

$$12{\cdot}3 - 5{\cdot}97 - 0{\cdot}04 = 6{\cdot}29 \text{ ev.}$$

This will be a quantum of wave-length 1740 Å. There will, however, be a very few photo-electrons on the wrong side of the threshold owing to the Maxwellian tail of the curve for 300° A. in the Quantum-Theory curves.

When the wave-length of incident radiation gets short enough for an X-ray, its quantum of energy is great enough to get the electron out of the lower energy-levels of the atoms in the material. In a general way we can regard thermions, and photo-electrons due to light softer than X-rays, as coming from the cloud of " free " electrons in the metal (not really free but rather lightly bound to the whole mass of metal as if it were an immense atom), and the photo-electrons excited by larger quanta from X-rays and γ-rays as coming from the inner levels of atoms themselves.

Photo-electric Cells

Though one is introduced to the photo-electric effect through Einstein's equation, $h\nu = \frac{1}{2}mv^2 + W_0$, so that one tends to think of the photo-electron as using some of the energy of the quantum $h\nu$ to get out of the surface, and the rest to give itself kinetic energy, one should think of this effect as only part of the whole P.E. effect. It is called the " outer " P.E. effect, and there is also an " inner " effect in insulators and semi-conductors by which extra electrons are detached from their atoms and given an amount of energy which enables them to take part in conduction. Such electrons do not leave the surface, and this is the distinction which must be kept clear. The outer P.E. effect expels electrons from the surface. The inner P.E. effect does not.

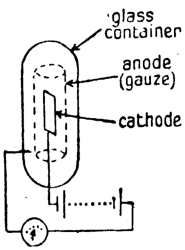


FIG. 286.

Photo-electric cells are instruments which use one or other of the photo-electric effects, and they are of several kinds. The emission type, Fig. 286, may be either evacuated or gas-filled. The vacuum type cells measure very small light intensities immediately and with great accuracy, for their emission is exactly proportional to the intensity of the incident light. Gas-filled cells (containing only a little gas) are about ten times as sensitive as vacuum cells, but are not accurate because the size of the gas effect is uncertain. This kind of cell, whose emitting surface is formed by evaporating

a very thin caesium layer on to an oxidized silver surface, tends to have a sensitivity extending far beyond the red into the infra-red.

The inner P.E. effect is used for two kinds of cell, called respectively " conductivity " and " rectifier " cells.

The conductivity cell, diagram Fig. 287, works by using a Wheatstone bridge arrangement to detect the alteration in resistance of the substance of the cell when light falls on its surface. Many materials, such as selenium, thallous sulphide, lead sulphide, have their electrical resistance lowered when light falls on the surface. These cells take a definite time to respond to the light fully, unlike the emission cells, which emit photo-electrons in less than 10^{-8} seconds after the quantum arrives (gaseous cells in about 10^{-6} seconds) and are

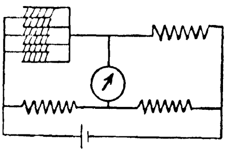


FIG. 287

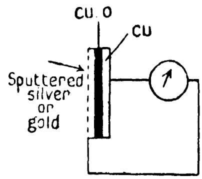


FIG. 288.

particularly red-sensitive--they have a maximum in the infra-red or sometimes the red—but are not accurate for quantitative work.

Rectifier cells (Fig. 288) use a sputtered silver or gold film on a copper oxide surface based on copper. A P.D. is developed when light falls on the sputtered surface. This kind of cell is strong and practical and gives a current of several milliamps, but the current is not truly proportional to the light intensity. Its advantage is that it is easy to construct and use, as it can easily be made up from an old metal rectifier and a galvanometer; but it will not follow rapid fluctuations of illumination.

The metal rectifier is a near relation. One side of a copper disc is oxidized by exposing it to a temperature of about 1000° C. The resulting disc of copper oxide on one side and copper on the other has about fifty times as high a resistance for currents flowing with the copper positive and the copper oxide negative than when it flows the other way; and the

disc thus makes a simple, strong, and efficient rectifier of alternating current.

Valves

In these devices, electrons pass from a hot cathode across an evacuated space to an anode. The flow of electrons may be modified by one or more subsidiary electrodes (" grids "), consisting of perforated metal plates or open-mesh wire spirals inserted between cathode and anode. The cathode in some types (including all the earlier ones) is a wire of pure tungsten, and has to be kept at about 2800° C. to get enough electron emission.

The cathodes of most newer types contain materials chosen for their low work-function so as to give enough emission at temperatures as low as 400° C.; sometimes the temperature is maintained by passing a heating current through the material of the cathode itself, and sometimes the heater coil is independent of, and electrically insulated from, the cathode, which is kept hot by radiation from the heater.

The enormous usefulness of valves depends on the almost complete absence of inertia of their moving parts, and hence their power of almost immediate response to minute and transient changes of potential or motions of charge. Ballistic galvanometers and quadrant electrometers can be made very sensitive, but can never be any good with transients because their inertia absolutely prevents a quick response. The valve also has the power (since the grid or third electrode was added) of greatly magnifying these transients within itself, without the complication of any extra apparatus. It is a relay without the ordinary relay's inertia. These are the two distinctive qualities that have made it so useful a tool in so many departments of physics.

The scientific and industrial applications of thermionics, in " valves," have become so widespread and so specialized that in a non-specialist text-book one can only try to get the general principles clear.

The Diode

The simplest kind of valve is the diode, or two-electrode valve. A triode with the grid and plate connected to each other will do for a diode in elementary experiments, which

the student should perform in order to understand the space-charge effect and Richardson's law for the variation of thermionic emission with temperature.

All the necessary experiments can be done with a simple circuit as in Fig. 289, in which F and P are the filament and plate of a valve, V a high-tension battery (two ordinary 120-volt wireless high-tension batteries in series will do for V),

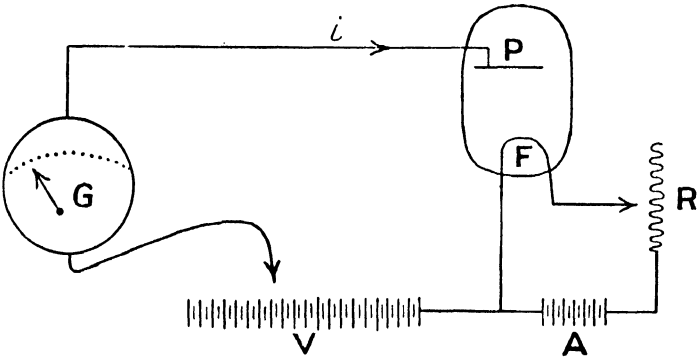


FIG. 289.

A an accumulator of 2, 4, or 6 volts, according to the filament used, and R a series resistance by which the current passing in F, and therefore the temperature of F, can be varied. G is a galvanometer of suitable sensitivity to carry the maximum current the valve is designed to produce. i is the current as measured by G.

The first experiment is designed to test the shape of the (V, i) curve before saturation is reached; that is to say, it tests the effect of the space-charge on the current flowing. A valve used for this purpose must be one with a pure tungsten filament.

To understand the space-charge effect, the student should set V_f, the filament voltage (measured by a voltmeter not shown in the circuit diagram), to the maximum allowed by the makers of the valve, to get full emission. Then he should plot i against V_p, the plate voltage, from quite small values up to the full value available. He should also arrange a reversing key to give negative values of V_p, because a small anode current flows when $V_p = 0$, and he should plot negative

values of V_p far enough to get the anode current to zero. The result will be a curve of the general shape of Fig. 290. This curve clearly divides itself into three parts: first, the part AB in which some current flows against an opposing plate voltage (obviously due to the initial energy of thermionic emission); second, the curved part BC from $V_p = 0$ till the curve begins to flatten out; third, the flat part CD when all the electrons emitted by the filament are getting to the plate, in spite of the space-charge, because V_p is big enough to overcome the space-charge. Why do I say the curved part represents the space-charge effect? For a reason that will now appear. If now log i is plotted against log V_p, a straight line, as shown (Fig. 291), is obtained.

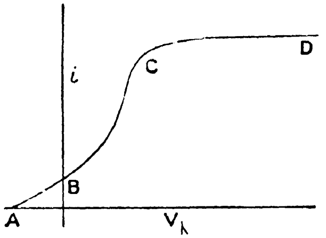


FIG. 290.

The slope of this straight line is $\frac{3}{2}$. It is evidence that over the curved part of the original curve $i \propto V^{\frac{3}{2}}$.

It can be shown that

$$i \propto (V_p + V_0)^{\frac{3}{2}}$$

if a space-charge operates, where V_0 is the average voltage through which electrons would drop to get the average energy with which they leave the filament. This agreement of experiment with theory is the evidence that the space-charge is operating.

To test Richardson's equation (p. 602) a series of such curves should be plotted for various values of V_f, the filament voltage.

The equation is roughly confirmed in its main principle if we get a straight line when we plot $\log\left(\frac{i}{T^2}\right)$ against $\frac{1}{T}$, where T is the absolute temperature of the filament. T is, however, awkward to find with the kind of apparatus one has handy for simple valve experiments, and it is possible to show from Stefan's Law (p. 613) that for fairly high temperatures $T^2 \propto V_f$. We can therefore get the required evidence by plotting $\log\left(\frac{i}{V_f}\right)$ against $\frac{1}{\sqrt{V_f}}$.

The argument runs as follows. The rate of supply of energy to the filament $\propto V_f^2$ (assuming the resistance nearly constant). By Stefan's Law, the radiation from a black body to surrounding at 0° A. is σT^4. The net loss of heat by radiation at temperature T_2 (Absolute) to surroundings at T_1 is strictly

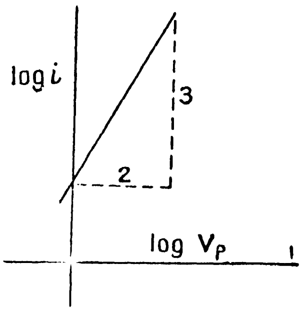


FIG. 291.

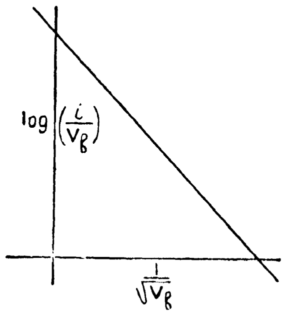


FIG. 292.

therefore $\sigma(T_2^4 - T_1^4)$, but in the range of temperatures in a thermionic valve (say 1200° to 300° Absolute) T_1 can be neglected, for if $T_2 = 4T_1$,

$$\frac{T_2^4 - T_1^4}{T_2^4 - 0} = \frac{256 - 1}{256} = \frac{255}{256},$$

giving a variation much smaller than that due to any of the other errors involved. Thus T^4 should vary as V_f^2:

$$T \propto \sqrt{V_f}.$$

As a diode only passes current in one direction, it can be used as a half-wave rectifier of alternating current, but when used in this way it is usually gas-filled to provide positive ions to neutralize the space-charge and so allow fairly heavy currents to pass.

The Triode

This valve was originally devised to deal with the space-charge difficulty. It has a grid, with an independently maintained potential, near the filament. It has characteristics which can be investigated by observing the variation of i_a with v_a for various values of v_g. The anode is connected through a milliameter to various tappings of an H.T. battery

of known voltage. The grid receives its potential from appropriate tappings of another H.T. battery and can be represented by a set of curves like those of Fig. 293.

These curves are interpreted by the use of three constants of the valve which we will now define. These are the following:

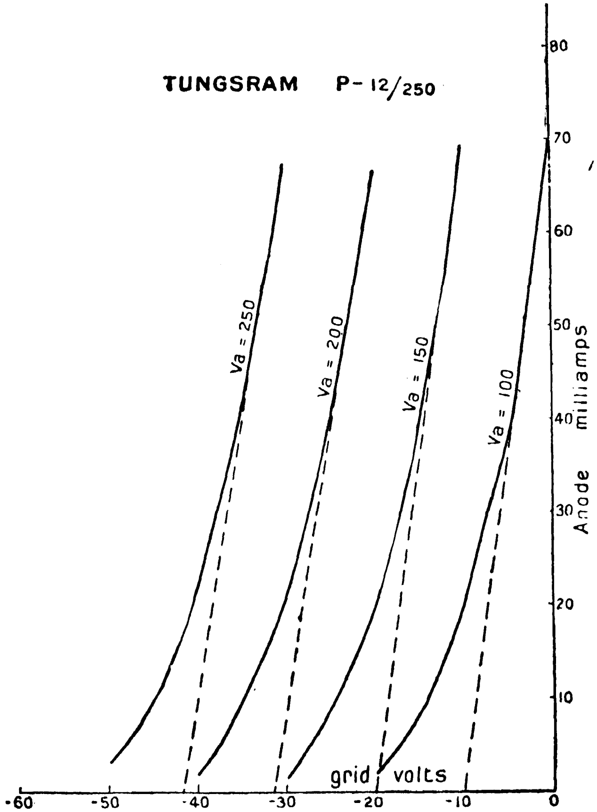


FIG. 293.

1. The amplification-factor, μ (pronounced *mew*, as in cat). This is the ratio of the increase of anode-voltage, causing a given increase of anode-current to the increase of grid-voltage, causing the same increase of anode-current. Thus

$$\mu = \frac{\partial i_a}{\partial v_g} \Big/ \frac{\partial i_a}{\partial v_a} = \frac{\partial v_a}{\partial v_g}.$$

This is only counted for the straight parts of the curves, which in the figure are produced down to the axis (in dotted lines) for convenience. Consider the region $V_g = 0 - 20$ volts and $v_a = 100 - 150$ volts. For the $v_a = 100$ and $V_a = 150$ curves, we can see that the slope

$$\frac{\partial i_a}{\partial v_g} = \frac{70}{10} = 7 \text{ milliamps. per volt.}$$

On the vertical line from $V_g = 10$ we can see by taking the intercept between the straight part of the curve for $v_a = 150$ to the (produced) straight line of the curve for $v_a = 100$, that for $V_g = 10$

$$\frac{\partial i_a}{\partial v_a} = \frac{70}{50} = 1{\cdot}4 \text{ milliamps. per volt.}$$

Thus the amplification-factor for this valve given by $\mu = \frac{7}{1{\cdot}4} = 5$ for this range.

2. The impedance, ρ; ρ is the Greek *r*, pronounced *rho* as in "rhododendron." It has nothing to do with the English *p* as in "pig."

The impedance is the reciprocal of $\frac{\partial i_a}{\partial v_a}$, v_g being kept constant. Thus

$$\rho = 1 \Big/ \frac{\partial i_a}{\partial v_a} = \frac{\partial v_a}{\partial i_a}.$$

In the curves in Fig. 293 it can be seen that

$$\rho = \frac{1}{0{\cdot}0014} = 713 \text{ ohms.}$$

3. The mutual conductance g_m is the rate of change of i_a with respect to v_g, v_a being kept constant. Thus

$$g_m = \frac{\partial i_a}{\partial v_g}.$$

Thus since

$$\rho = \frac{\partial v_a}{\partial i_a} \text{ and } \mu = \frac{\partial v_a}{\partial v_g},$$

we have

$$g_m = \frac{\partial i_a}{\partial v_g} = \frac{\partial v_a}{\partial v_g} \Big/ \frac{\partial v_a}{\partial i_a} = \frac{\mu}{\rho},$$

similarly

$$\mu = \rho . g_m$$

$$\rho = \frac{\mu}{g_m}.$$

Thus for the curves in Fig. 293

$$g_m = \frac{\mu}{\rho} = \frac{5}{713} = 0{\cdot}007 \text{ amp. per volt}$$

$$= 7 \text{ milliamps. per volt.}$$

Let us now consider the range $v_g = 20 - 40$ volts and $v_a = 200 - 250$ volts.

For $v_a = 200$ and 250 volts respectively

$$g_m = \frac{\partial i_a}{\partial v_g} = \frac{59}{10}, \frac{55}{10}, \text{ respectively; average } 5{\cdot}7 \text{ ma./volt}$$

$$\frac{\partial i_a}{\partial v_a} = \frac{57}{50} = 1{\cdot}14 \text{ ma./volt.}$$ [Intercept on the ($v_g = 30$) ordinate.]

Thus

$$\mu = \frac{5{\cdot}7}{1{\cdot}14} = 5{\cdot}0$$

$$\rho = \frac{50}{{\cdot}057} = 877 \text{ ohms.}$$

The makers give, for this valve, $\mu = 5$, $\rho = 850$ ohms, $g_m = 6{\cdot}0$ ma./volt.

Tabulating these results we get:

Source	μ	ρ *ohms*	G_m. *ma. per volt*
$V_g = -10$ $V_a = 125$	5	713	7·0
$V_g = -30$ $V_a = 225$	5	877	5·8
Makers	5	850	6·0

Voltage Amplification by the Triode

When a triode is used as an amplifier, this is done by making it increase the amplitude of an applied alternating E.M.F.

The alternating E.M.F. is applied between grid and filament, and a magnified form of it, $R\delta i_a$, is produced across the

resistance R connected between the anode and the source of anode potential.

We have two separate partial causes of increase of anode current $\frac{\partial i_a}{\partial v_g}\left(= g_m\right)$ and $\frac{\partial i_a}{\partial v_a}\left(= \frac{1}{\rho}\right)$.

Thus the complete increase of current δi_a is obtained by adding the effects due to these separate causes. Thus

$$\delta i_a = g_m \delta v_g + \frac{1}{\rho}\delta v_a.$$

Now the voltage applied to the plate is reduced when i_a is increased, because the lost volts across R get bigger. Thus

$$\delta v_a = -\mathrm{R}\delta i_a.$$

Substituting, we get

$$\delta i_a = g_m \delta v_g - \frac{\mathrm{R}}{\rho}\delta i_a$$

$$\delta i_a\left(1 + \frac{\mathrm{R}}{\rho}\right) = g_m \delta v_g.$$

Now the voltage amplification is given by

$$\frac{\text{Voltage change across R}}{\delta v_g} = \frac{\mathrm{R}\delta i_a}{\delta v_g} = \frac{\mathrm{R}g_m}{\left(1 + \frac{\mathrm{R}}{\rho}\right)} = \frac{\mathrm{R}\rho . g_m}{\mathrm{R} + \rho}.$$

But we have shown that $\rho . g_m = \mu$. So we have

$$\text{Voltage amplification} = \frac{\mu \mathrm{R}}{\mathrm{R} + \rho}$$

which shows that it gets bigger as R is made bigger, but can never be as large as μ.

Note that since v_a is of quite high voltage, and normally supplies a current of several milliamps., whereas v_g is normally a much smaller voltage supplying a very much smaller current, there is not merely voltage-amplification (as in a transformer) but also power amplification.

Many other types of valve are used, in which the principles outlined above are elaborated. Tetrodes having a second grid and pentodes having a third grid give extensions of working conditions, and especially much greater amplification, but the reader who wants to go further should consult the specialist books on this immense subject.

Accelerators

For producing the high voltages and immense particle energies required in modern nuclear physics, a number of special machines have been devised.

The Cyclotron

Imagine two semicircular magnetized discs AB placed as in Figs. 294 and 295, one above the other, with a gap between

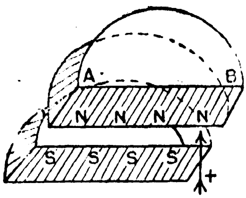


FIG. 294.

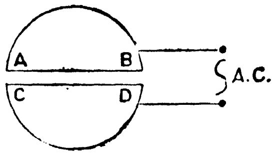


FIG. 295.

them. A positive ion projected between them along the line the arrow would (with the right velocity) describe a circle in anti-clockwise direction, seen from the top. (Fleming's Left-hand Rule.)

Now imagine another pair of similar discs CD opposite, and look down on the set of discs from the top. If the two pairs of discs are separated by a vertical gap, and connected to a source of A.C. supply, an alternating electric field would act between AB and CD. If the speed of the alternations and the relative strengths of the electric and magnetic fields were timed just right, the circulating positive ion could get a push from the electric field as it crossed the gap each time, going either way. It would thus be accelerated, and if it were first projected not far from the centre it would describe a spiral, the radius of curvature of its path gradually increasing with its speed, until it could finally be projected from the outer circumference of the "Dees" with great speed. This is the arrangement of the cyclotron. Clearly it cannot work unless the time of revolution is independent of the velocity, but it can easily be shown to be so.

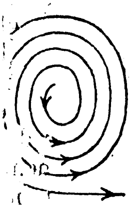


FIG. 296.

discovered by Urey, Brickwedde, and Murphy in 1932 as a result of a special search. When the heavier isotopes of oxygen were discovered, the whole atomic weight scheme based upon the assumption O = 16·0000 went wrong. Instead, on the assumption that $_8O^{16}$ = 16·0000, it followe that the chemical weight (*i.e.* the average weight includin all isotopes) on the O^{16} = 16·0000 scale was 16·0044. Up to ow the mass-spectrograph and the chemical weight of hydro en had agreed nicely at 1·00778 and 1·00777 respectively. Applying the new factor $\frac{16{\cdot}0044}{16{\cdot}0000}$ made the chemical weight too large, and caused the search for a heavy isotope. (Actually both the figures 1·00778 and 1·00777 also proved to need revision, so that a false clue led to a true conviction.) The 1948 value of the $_1D^2$ mass is 2·01473 on the O^{16} = 16·0000 scale. This includes the electron and is the mass of th atom of heavy hydrogen, also called deuterium. The $_1D^2$ nucleus alone gives 2·01418.

Cosmic Rays

It has been known since about 1850 that the air is alwa slightly conducting in normal circumstances, or (ir words) that normal air is always slightly ionized. When matter was investigated round about 1900, it was thou that traces of radioactive elements in earth or air might the sufficient cause. It was not known then that the effect stronger in the upper atmosphere, or that it is diminish almost to vanishing point at the bottom of a mine or lake. We do now know for certain from these experim results that rays with several components, some of them m penetrating than any other rays we know, do rain down the earth from somewhere outside. Many facts have been established about them from experimental evidence they are called cosmic rays because they come fro directions, not from any special place. They come from cosmos. The primary cosmic rays, as they arrive at the of the earth's atmosphere, seem to contain protons and particles, with a small proportion of heavier nuclei with ato numbers between 6 and 26.

The heavier nuclei are nearly all stopped by collisi

before they get within twelve miles of the earth's surface. The surviving protons are so much thinned out on the way down that for every collision at sea level there are about 3000 at an altitude of fourteen miles (according to pilot-balloon observations). The most energetic primaries are hundreds of times as energetic as anything we can produce artificially on earth at present (1949).

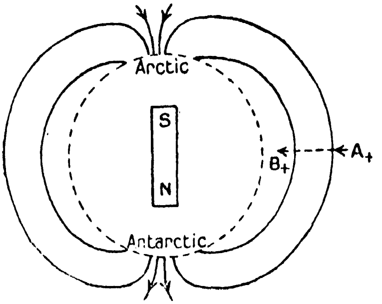


FIG. 298.

As a result of collisions in the atmosphere, secondary electrons, photons, neutrons, neutrinos, mesons, and positrons (see below) are produced. There is at present (1949) a difference of opinion as to whether the primary rays contain electrons as well as positive nuclei, and to explain this difference of opinion we must consider the effect of the earth's magnetic field upon incoming charged particles.

In Fig. 298 the dotted circle represents the earth's surface, the bar-magnet a totally imaginary cause for the earth's field, the hard curved lines the direction of the field outside the earth. The arrow A_+ B_+ represents a $+$-charged particle moving from outside to the equator. By the left-hand rule it will experience a force acting downwards into the paper; *i.e.* in an easterly direction with respect to the earth's surface. An observer on earth who detected this particle would therefore see it coming from the west. A preponderance at the earth's surface of particles from the west means a preponderance of positively charged particles of high penetration. This is observed. But the preponderance keeps on getting less as we go up from the earth's surface; and it therefore seems likely that a set of particles of less penetration and negative charge is also arriving. These are presumably electrons. It seems likely, on general grounds, that the average numbers of positively and negatively charged particles arriving would preserve an even balance, for the evidence from ionization indicates that about $8{\cdot}0 \times 10^{17}$ primary particles per second reach the earth from outside. If these

were all $+^{ve}$, they would carry a current of about 0·13 amps. (which does not seem much), but it would increase the electric potential of the earth by about 180 volts per second. The average energy of the particles is about $7{\cdot}7 \times 10^9$ ev, so that the power involved is about a million kilowatts. The question is discussed in the January 1949 number of *Reviews of Modern Physics*, which is fully devoted to cosmic rays and mesons.

The earth's magnetic field has a further effect which is equally easy to understand. Particles may be considered to approach the earth from all directions. There is no evidence of a favoured direction. The deflection of a moving charged particle in a magnetic field depends on the angle which the direction of motion of the particle makes with the field. It is zero at 0°, a maximum at 90°. Thus particles moving along the lines of the magnetic field can get straight to earth undeflected, provided they have enough energy to penetrate the atmosphere. Thus, of particles with just enough energy to reach earth, the least energetic will be those getting to the poles. There will, at each latitude, be a threshold energy which will just get the particle to earth, and this will increase gradually from pole to equator. The higher the latitude, therefore, the more particles will arrive at the surface. The energy required to get a particle just to the earth's surface seems to range from about 8×10^9 ev at the equator to about 2×10^9 ev at the poles.

Attempts have been made to trace variations in intensity of the cosmic rays as a whole as they reach the earth, for the sake of the extremely interesting possibilities one might explore in this way. For example, owing to revolution of our galaxy as a whole, the solar system, and the earth with it, have a velocity of translation of about 250 miles per second in the direction of the earth's North pole. This motion might cause greater cosmic-ray intensity in the northern than in the southern hemisphere, and for some time this effect was thought to show; but the divergence is no longer regarded as being larger than experimental error, so that the existence of the effect can be neither confirmed nor denied. We do know that cosmic rays do not come from the sun, but we do not know whether they come from within our own galaxy or from outside it. Such daily and seasonal changes as we do observe

seem to be due to local and minor effects, such as magnetic storms, the sun's magnetic field, and variations in condition of our own upper atmosphere which affect the production of secondary effects there. There is evidence of sudden increases of intensity of cosmic radiation, beginning about an hour after the appearance of a large solar flare, reaching a maximum rapidly, and dying down in about a day.

There is a remarkable distribution of energy among primary particles. Few particles seem to arrive at the top of the earth's atmosphere with an energy of less than several MeV. The only satisfactory way known to account for this is to assume that most of the less energetic particles are captured by the magnetic field of the sun. The planets may get some too. This assumption is consistent with the view that the primary cosmic rays come from parts of the cosmos outside the solar system.

Mesons

The meson was inferred theoretically by the Japanese physicist Yukawa in 1935. Its discovery in a Wilson cloud-chamber was announced the next year by Anderson, four years after he discovered the positron. The name meson means middle, and was given to the particle because its mass is between the proton's mass and the electron's.

Yukawa considered the binding forces within the nucleus as analogous to those which hold atoms together to form a molecule. Since chemical bonds are associated with the sharing of electrons, Yukawa supposed such binding-by-sharing to be associated with the mass of the shared particles, and he devised an equation connecting the mass of the shared particles with the distances between the parts bound. The actual coefficient he used was known as η, which was the reciprocal of the distance over which the binding forces have to act. His equation for the rest-mass of the exchanging particle or particles was given by

$$m_0 = \frac{h\eta}{2\pi c}.$$

He then asked himself what exchange particles could bind protons and neutrons together in a nucleus, and he took the effective distance between them as between 2×10^{-13} cm.

and 3×10^{-13} cm., or thereabouts. If we substitute values from pp. 660–1 we get from 193 to 129 electron rest-masses; but as the π-meson, or pion—now recognized as what Yukawa inferred—has a mass of 273, it seems that about $1{\cdot}4 \times 10^{-13}$ cm. would have been a better estimate for the effective distance. Yukawa further deduced that his particle would have a very short life, a microsecond or less. It would then disintegrate (he thought) into an electron, positive or negative, and a neutrino. Yukawa received the Nobel Prize for Physics in 1949.

A meson, thought at first to be Yukawa's, was detected by Anderson and Neddermayer in 1936, in a Wilson cloud-chamber; the curvature and ionization of this particle's track were about right for Yukawa's upper estimate of 192, and it therefore could not possibly be an electron or proton. This meson's disintegration was photographed by E. J. Williams and E. Pickup of the University College of Wales, Aberystwyth, in 1938; and the photograph showed the meson vanishing and the electron moving off in a different direction—clear evidence that an unobserved neutral particle was a second product. Yukawa's predictions seemed to be closely verified. However, the facts accumulated to show that the observed meson was not Yukawa's, but one of its disintegration products, now known as a μ-meson (mu-meson).

Yukawa's π-meson is produced in the upper atmosphere by the encounter of a cosmic-ray nucleon with an atmospheric nucleon. After its short life, much shorter than Yukawa expected, it decays into a μ-meson and a neutral particle. The μ-meson decays, with about the life Yukawa estimated for the π-meson, into an electron and two neutral particles. It is immensely penetrating, and is observed not only at the earth's surface but also below it.

Since 1947, knowledge about mesons has proliferated at an astonishing speed, largely through the work of Prof. C. F. Powell and his team at Bristol, with their technique of recording atomic events at very great heights in special emulsions on photographic film sent up attached to special balloons. These balloons rise to heights of 100,000 feet or more before returning to earth, and are generally recovered with their film undamaged. By this technique it is possible to record events involving mesons, and also other events which could not

occur at low levels. The films can then be scanned at leisure in the laboratory by highly skilled teams.

The two main kinds of mesons now (1955) known are called L-mesons (mass between 200 and 300) and K-mesons (mass between 900 and 1,100). There are also the heavy hyperons (mass between 2,100 and 2,400) which can probably be regarded as temporarily excited nucleons.

The facts about these particles are summarized in the following (1955) tables:

Name	*Symbol & Charge*	*Mass*	*Half-life*	*Decay*
		LIGHTEST GROUP. L-MESONS		
Mu-meson or muon	$\mu^{\pm}$	$207{\cdot}0 \pm 0{\cdot}4$	$(2{\cdot}22 \pm 0{\cdot}02) \times 10^{-6}$ sec.	$\mu^{\pm} \rightarrow e^{\pm} + 2\nu$
Pi-meson or pion	$\pi^{\pm}$	$273{\cdot}0 \pm 0{\cdot}4$	$(2{\cdot}53 \pm 0{\cdot}1) \times 10^{-8}$ sec.	$\pi^{\pm} \rightarrow \mu^{\pm} + \nu$
Neutral Pi-meson	π^0	264 ± 1	3×10^{-14} sec.	$\pi_0 \rightarrow 2\gamma$; $\pi_0 \rightarrow \gamma + e^+ + e^-$
		MEDIUM GROUP. K-MESONS		
Tau-meson or tauon	$T^{\pm}$	965 ± 3	3×10^{-9} sec. approx.	$\tau^{\pm} \rightarrow \pi^+ + \pi^- + \pi^{\pm}$
Theta-meson	θ^0	970 ± 10	10^{-10} sec. approx.	$\theta^0 \rightarrow \pi^+ + \pi^- + \pi^0$?
Kappa-meson	$K^{\pm}$	about 1000	? 10^{-9} sec.	$\kappa^{\pm} \rightarrow \mu^{\pm} + 2\nu$?
Chi-meson	$\chi^{\pm}$	about 1000	?	$\chi^{\pm} \rightarrow \pi^{\pm} + \gamma$?
		HEAVY GROUP. HYPERONS		
Omega	$\Omega^{\pm}$	about 2300	?	Ω? $\rightarrow n^0 + \pi^{\pm}$
Lambda	Λ^0	2181 ± 1	3×10^{-10} sec. approx	$\Lambda^0 \rightarrow p^+ + \pi^-$

The positron

The positron has a charge and mass believed to be identical with those of a negative electron, or negatron, except that the charge is positive. Positrons are no more subject to exponential decay than negatrons; but nevertheless a positron, in the presence of matter of ordinary density, is annihilated within about 10 seconds by combination with a negatron, and the mass of both particles is converted to -radiation. As the number of free positrons in our cosmos is very small compared with the number of free negatrons, negatrons seem stable and positrons short-lived. Protons seem to be the only fundamental particles which are permanent.

The actual discovery of the positron by Anderson in 1932 was due to a lucky accident. Particles resulting from cosmic rays had such immense energy, compared with any energies previously known, that a very large Wilson cloud-chamber, with an immensely powerful magnet (field about 24,000 oersteds) was needed to deal with them. Electrons with a few MeV energy described minute spirals in such a field, which could produce observable curvature in electrons of 6000 MeV. In such a cloud-chamber paired tracks were observed, as of particles of similar mass but opposite charge. Confirmation was obtained by allowing the supposed positive electron to penetrate a lead plate 6 mm. thick. Curvature showed an energy of 63 MeV on one side, and 23 MeV on the other, and this showed conclusively which way the particle was going. The direction of bending conclusively proved the charge positive. It could not be a proton. No proton having the amount of energy associated with the ionization shown by this particle could possibly have penetrated 6 mm. of lead. The only answer was the positron, and Anderson was awarded the Nobel prize. Further experiment showed that the specific ionization produced and the measured value of $\frac{e}{m}$ were the same (within the limits of experimental error) as those of the electron.

Why had the positron not been found before? The answer seems to be partly in its short life before its mutual annihilation with an electron, and partly in the fact that either as a cosmic ray product of immense energy or as a product of artificially radioactive transformations it hardly could have been observed before.

Before the positron was found, a positive particle probably lighter than the proton had been predicted by Dirac on theoretical grounds. His very remarkable line of thought was somewhat as follows.

(1) The Relativity Theory gives a quadratic general equation for the total energy of a moving particle. If p is the momentum, this equation is

$$E^2 = c^2p^2 + m_0^2c^4.$$

The reader can obtain this equation by simple algebra if he eliminates v from the two equations for the momentum and

energy of real moving particles. The first of the two equations can only refer to real particles because E is positive.

$$E = \frac{m_0c^2}{\sqrt{1 - \beta^2}}, \qquad \text{where } \beta = \frac{v}{c}$$

$$p = \frac{m_0v}{\sqrt{1 - \beta^2}}.$$

The positive value for E is not the only one possible, for the quadratic in E gives two solutions, one positive and one negative.

How can the total energy of a particle be negative? or what meaning, if any, can negative energy have?

(2) There is a fundamental difficulty in connection with the interaction between matter and radiation. According to wave-mechanical theory, a harmonic oscillator still has a finite amplitude of vibration in its lowest energy state. Space as the vehicle of radiation must be regarded as containing such harmonic oscillators. Even in "empty" space, with no quanta or particles present, there is thus what is called "zero-point fluctuation" of the energy, and the root-mean-square values of the field-strength are not zero. We must therefore regard "empty" space as having a structure.

What is this structure?

Two incomprehensibles are sometimes less intractable than one. They may help to explain each other; and these two problems, of the structure of empty space and the nature of negative energy, seem to have done this. Pauli's exclusion principle states that no energy state in a system can be occupied by more than one electron. Dirac applied this principle to the negative energy states, and assumed that each energy state whether positive or negative, could be "occupied" by an electron, provided only that its energy, whether positive or negative, could have a value consistent with the energy equation

$$E^2 = c^2p^2 + m_0^2c^4.$$

This requirement limits the energy states to those states not less positive than $+ m_0c^2$, and not less negative than $- m_0c^2$, since these two values are the smallest possible numerical values for E, obtained when $p = 0$ and the particle is at rest. No values between $+ m_0c^2$ and $- m_0c^2$ are possible for a particle with a rest-mass m_0 (though they might be for a particle of less mass, such as a neutrino).

Thus we can regard ourselves as surrounded by observable electrons with positive energy not less than $+ m_0c^2$, and also by unobservable electrons with negative energy, numerically not less than $- m_0c^2$. By the exclusion principle, each of these negative energy states can only contain one of these "ghost" electrons, and the "ghost" can only materialize by receiving a quantum of energy large enough to cancel out its negative energy and give it positive energy not less than m_0c^2. (This "materializing" quantum must therefore in any case be as large as $2m_0c^2$ (1·022 MeV),[1] and will in general be much larger.) When this happens, the negative energy state vacated by the "ghost" electron, now materialized, will be "empty"—a hole in space. The mathematical description for this "hole," due to the "absence" of a ghost negative electron, turns out to be also an expression for the presence of a positive electron. Dirac postulated that this expression, which had emerged from the mathematical theory with a double interpretation, might in respect of one of its interpretations be found to refer to an observable particle. This theory seems likely to be part of the interim stage of a much more general theory covering all fields and particles.

Pair Production, Showers, Stars

On the evidence it seems that the following are among the possible effects of the arrival of primary cosmic rays at the top of the atmosphere.

1. Atomic nuclei are disrupted and mesons are projected with very high energy. These form the most penetrating radiation, a radiation so penetrating that (for its hardest component) the absorption coefficient is about 0·0025 per metre of water. This radiation can still be observed after penetrating 1400 metres of water or its equivalent, 124 metres of lead. Its intensity amounts to only about one part in 20,000 of that of the primary radiation. The absorption coefficient appears to decrease (*i.e.* the rays appear to get harder) as one goes down, but really, of course, the softer radiation is gradually getting more and more filtered out until only the very hardest is left.

[1] For $2m_0c^2 = 2 \times 9 \times 10^{-28} \times 9 \times 10^{20}$ ergs, nearly $= 1{\cdot}6 \times 10^{-6}$ ergs $= 1$ MeV, nearly.

2. The impact of one of the primary particles on a nucleus causes a photon of more than 1 MeV energy to be emitted. A photon whose wave-length is less than the wave-length for which $h\nu = 1{\cdot}6 \times 10^{-6}$ ergs, or $\frac{hc}{\lambda} = 1{\cdot}6 \times 10^{-6}$ ergs.
Thus

$$\lambda = \frac{6{\cdot}6 \times 10^{-27} \times 3 \times 10^{10}}{1{\cdot}6 \times 10^{-6}} = 1{\cdot}24 \times 10^{-10}\text{ cm} = 0{\cdot}0124\ \text{Å}.$$

As we have seen (p. 633), the rest-energy of an electron or positron is close on 0·5 MeV, so that a photon of 1 MeV or over should be able to create a pair. γ-rays, such as the 2·65 MeV γ-rays from Th.C″, can and do do this, and an event of this kind is thoroughly consistent with Dirac's theory of the positron, as explained in the last section.

3. With primary particles of high energy, the pairs or mesons produced by secondary γ-rays may have enough energy to produce more γ-rays by impact which themselves may proceed to produce more pairs with enough energy in the primaries, this may happen again and again, the total number of particles being doubled at each collision with an atomic nucleus. Such a collection of particles, all originating from a single one of very high energy, is known as a shower, and can be observed in process of formation in a Wilson cloud-chamber containing successive thin lead screens to provide matter for collisions to produce γ-rays.

It is from such showers that evidence comes of the existence of primaries of very great energy. There may be thousands of particles in a single shower. Auger in 1938 made observations which indicated that a single shower might be 10 metres wide and contain over a million particles, so that the original particle apparently must have had an energy of about 10^{15} ev, 10^{9} MeV. Some observations by Lapp in 1943 suggest that occasionally the energy of the primary particle is as much as 10^{17} ev or 10^{11} MeV. If a helium atom of mass 4 mass units produces primaries of energy 2×10^{6} MeV (actually about 1·9 mass units), as we have seen, then the 10^{14} MeV primary particle presumably originated as one of a pair from something of mass about 2×10^{8} mass units. A heavy atom like lead weighs about 200 mass units, so the suggestion is that something of the mass of a million lead atoms was dissipated to produce one pair of primary particles. This extremely odd

conclusion seems difficult to account for on any theory whatever. Special kinds of shower called " bursts," probably due to the collision of shower particles with the walls of the chamber, also occur, sometimes small showers entirely composed of mesons can penetrate up to 15 cm. of lead. These occur specially often at higher altitudes, and have the curious property of not themselves generating further showers in a lead plate. There is another rare but interesting kind of shower called a " star." The total energy of such showers is relatively low—with apparently a total of under 10^6 MeV for the whole star and up to 100 MeV for single particles. These particles appear to be heavy ones, making short thick tracks. Some seem to be protons and there are some α-particles. The whole star indicates some sort of nuclear break-up caused by a cosmic-ray particle.

Pair production, and its converse the production of photons through the annihilation of particles, seem to be of fundamental importance to our understanding of the material universe. Pair production has been described at the end of the positron section. Dirac's theory gives us a way of thinking about both production and annihilation. When there is mutual annihilation of an electron and positron (happening, like pair production, only in a very strong field) the explanation on Dirac's theory, so far as it can be put into words, is somewhat as follows. An observable electron has positive energy which cannot be less than its rest energy, m_0c^2 (0·511 MeV). It can, however, lose more energy than this, provided only that it loses enough to enable it to occupy one of the " holes," or negative-energy states, we have discussed. The energy of one of these states must be at least as negative as $-m_0c^2$, or it could not be a possible solution for the equation $E^2 = p^2c^2 + m_0{}^2c^4$. To be annihilated with a positron, therefore, an electron must lose not less energy than $m_0c^2-(-m_0c^2)$, or $2m_0c^2$. This energy may reappear either as two equal photons going in opposite directions, or as a single photon. In the latter case, the compensating momentum may be received by the nucleus in whose field the event has happened. When the electron occupies the " hole," or negative-energy state, it becomes unobservable, and so does the positron which was the outward and observable sign of the empty " hole "; and so both are said to be " annihilated." Millikan has made

the exceedingly interesting suggestion that the annihilation of heavier nuclei in the cosmos may be the origin, after a suitable series of exchanges, of the primary cosmic rays of very high energy. He gives an easily understood account of his work in this field in *Reviews of Modern Physics*, January 1949. The theory is not at present, 1949, accepted by the majority of physicists. He suggests that to find the origin of something of 1 thousand MeV energy, one should direct one's search to an entity of 2 thousand MeV energy or 2 mass units. Thus the results of dissipating $_1H^1$, $_2He^4$, $_6C^{12}$, $_7N^{14}$, $_8O^{16}$, $_{14}Si^{28}$, and $_{20}Ca^{40}$ would be particles of 0·5, 1·9, 5·6, 6·6, 7·5, 13·2, 21·3 thousand MeV. Millikan has suggested that the evidence does indicate something like line velocity spectra of cosmic rays with these energies, or at least some of them. This theory is no help for particles of 10^{14} MeV energy.

Nuclear Stability and the Packing Fraction

We will now consider the effects of the knowledge and power of action we have gained by the present refinement of our methods of measuring atomic masses, and by the discovery of the neutron; for perhaps these are the discoveries which at present (1949) are doing most to change the world materially for better or worse.

We may regard Prout's hypothesis as finally confirmed. On the atomic weight scale on which $O^{16} = 16·0000$, every nucleus found up to 1949 departs from the whole number rule by the third place of decimals, but only two (U and Th) in the first place of decimals. Every atomic mass from H^1 to F^{19} has an excess (except O^{16}), every mass from Ne^{20} to Gd^{157} has a defect.

" To have an excess " means " to have a mass greater than the nearest whole number," and " to have a defect " means "to have a mass less than the nearest whole number." Thus $_1H^1$, 1·0077, has an excess, and $_{25}Mn^{55}$, 54·965, has a defect. From Gd^{158} to Os^{189} measurement is not yet accurate enough to tell. From Os^{190} to U^{238} there is an excess. The question of excess or defect is not important, for it depends on the more or less accidental choice of O^{16} as a standard. If Mn^{55}, the atom with the biggest relative defect, had been chosen, every atom would have had an excess, except for Mn^{55}. If H^1 had been chosen (and I cannot help thinking this would

have been best) every atom would have had a defect, but Prout's Law would have appeared to fail.

What does the variation from Prout's Law signify?

The answer to this question is quite clear. On the relativity theory all energy, whether kinetic, potential, or in any other form, has mass. A nucleus is a system of forces. The lower its mass for a given number of particles of known rest-mass, the lower its potential energy, and the more stable the system. In a general way we can see how this works by plotting a graph of mass excess against atomic weight for the lighter elements. In Fig. 300 we see that (except for Li^8) there is a marked drop for He^4, C^{12}, O^{16}, Ne^{20}, all of which have atomic weights which are multiples of 4.

Now the He^4 nucleus is probably the most stable of all, or it would not stand the racket of being an α-particle. It is reasonable to suppose that light atoms with a mass number which is a multiple of 4 should be specially stable and have a lower mass compared with their whole-number mass. This seems to be borne out for He^4, N^{12}, O^{16}, Ne^{20}, Mg^{24}, but not for either of the rare isotopes Li^8 and Be^8, both artificially produced, both radioactive, and one of them (Li^8, which does show a mass-drop comparable with that of He^4 and C^{12}) able to release 12 MeV on disintegration with a half-life of 0·88 seconds.

Besides the mass excess and defect curves, we should consider what Aston called the packing-fraction curves. These are obtained by plotting against the atomic weight (see Fig. 299) the mass-drop per mass-unit, obtained by dividing the mass-defect or mass-excess by the atomic weight. Thus U^{238} has a measured atomic weight of 238·14. Its packing fraction

$$= \frac{0{\cdot}14}{238} = 5{\cdot}6 \text{ parts per } 10{,}000.$$

The curve shown in Fig. 299, showing a representative selection of elements, has some points worth comment. The beginning of it shows two branches, one for the specially stable He^4, N^{12}, O^{16} (shown with crosses, as are also Ne^{20}, Mg^{24}, Si^{28}), and the other for other elements. The general curve is high for light and heavy elements and low in the middle, with its minimum in the neighbourhood of iron and cobalt. In a general way this seems to make sense. One

may reasonably suppose that light nuclei have not got enough particles to enable them to lower their potential energy by efficient packing, like the toothbrush-and-comb week-ender who arrives with bulging pockets. Energy-reduction per unit packed increases (or efficiency of packing increases) till the

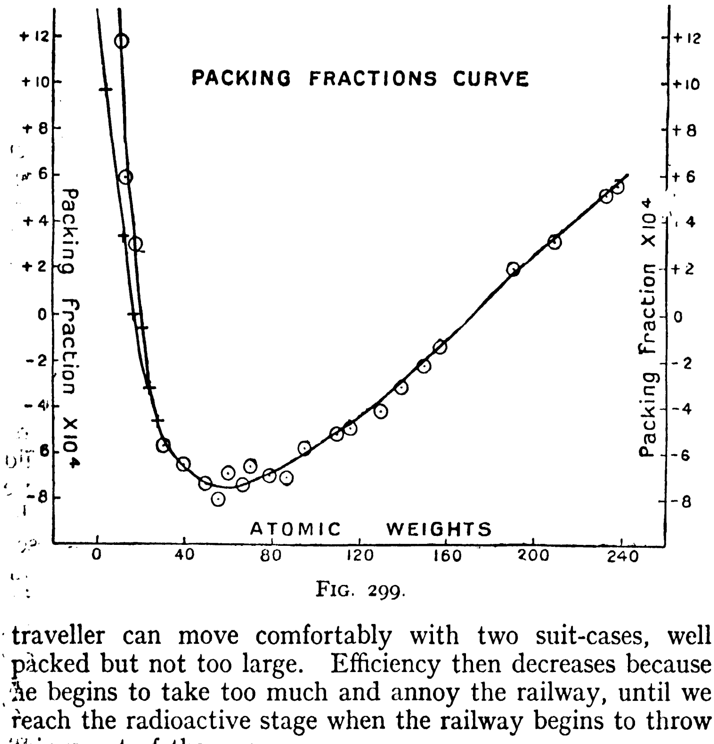


FIG. 299.

traveller can move comfortably with two suit-cases, well packed but not too large. Efficiency then decreases because he begins to take too much and annoy the railway, until we reach the radioactive stage when the railway begins to throw things out of the van.

We now have a picture of a nucleus which makes sense of artificial creation of new isotopes and even new elements, he radioactive and some not; and also of the development available atomic energy. We can see the nucleus from helium (or from D^2), upwards, as an immensely complicated structure of protons and neutrons with energy-levels belonging to the structure as a whole. Alternatively we may think of the proton and neutron as alternative states of the same particle (for which Rutherford suggested the name nucleon).

theory. There is more common sense and less difficult philosophy in it (at any rate for a beginner) than one might be tempted to suppose by reading what philosophers have written about it; and anyway a physicist should always have his feet on the firm ground of experimental evidence and simple numerical relations; and in connection with the Uncertainty Principle it is quite easy to stand on this firm ground.

To begin with, the word "uncertainty" is rather unfortunate, and so is the other commonly used word, "indeterminacy," because both suggest a certain woolliness, which is the exact opposite of what we want. Heisenberg's word is *Unbestimmtheit*, which might also have been called "indefiniteness," which seems quite a good word to me. As a rough analogy, imagine yourself estimating as a percentage the unreliability of witnesses in a law court. They are testifying to a past event, since proved to be associated with a crime, but not specially regarded by them at the time it happened. A witness who testifies that the event happened between 11.00 a.m. and 12.00 noon you might rate as 0% unreliable. A witness who testified that it happened between 11 hrs. 23 min. 17·7 sec. and 11 hrs. 23 min. 17·8 sec. you would unhesitatingly rate as 100% unreliable. It would be quite sensible to say that the product of your estimate of the percentage unreliability and the wideness of the limits chosen by the witness would be roughly constant. This is only an analogy, and the reader should not mix up Heisenberg's indefiniteness, claimed as really in the nature of things, with inexactness of statement. The law-court analogy is only intended to bring out the fact that absence of definiteness in a statement has nothing whatever to do with absence of truthfulness. On the contrary the story illustrates my preference for "indefiniteness" as a translation of Heisenberg's word. The uncertainty lies with the witness who chooses narrow time-limits, not with the cautious witness. Reliability goes with modesty of statement, and Heisenberg and the modern physicists are more modest of statement than was Laplace when he claimed that a sufficiently competent mathematician, given the mass and motion-vectors of every particle of the universe accurately, could foretell the future exactly. When you read that the uncertainty principle has destroyed

causality, I think you should not be too excited about this, but take it that from the point of view of the theoretical physicist Laplace's view cannot be sustained. As Born says, the law of causation, *in the sense of classical physics*, has lost its validity. Nor is this discovery new; it is merely newly realized. If you send a single pulse of light, one wave-length long, through a spectrometer, the spectrometer will analyse it out into a wide range of wave-lengths, even if it is a sinusoidal pulse. You cannot get monochromatic light, *i.e.* light whose wave-length can be stated within very close limits, without having it in the form of a long train of waves.

What is new in Heisenberg's discovery is its exact mathematical (or rather, numerical) statement, and its association with Planck's quantum h. What Heisenberg says is that in any small system, such as a photon or electron (he does not himself make this restriction, but for practical purposes we may take him to mean it), there is always some indefiniteness. This indefiniteness applies to both the position and the momentum, when the two are measured simultaneously.

The photon and the electron must now both be regarded as having position and momentum. Since the electron has been shown (theoretically by de Broglie and experimentally by G. P. Thomson and others) to have wave-length, they both have wave-length. If x is the position-co-ordinate and your best estimate is that its value is between x and $(x + \delta x)$, and p is the momentum and your best estimate is that it lies between p and $(p + \delta p)$, then

$$\delta x \times \delta p \doteq h.$$

Note (a) that the dimensions are right (each side has dimensions ML^2T^{-1}), and (b) that an *approximate* "equals" sign is used.

It can, for example, be shown (but the true proof is beyond the scope of this book, and a shortened form, though apparently intelligible, might be misleading—see Born's *Atomic Physics*, Appendix XII)—that if we build up a wave-packet from separate wave-trains we need a definite frequency-range in these wave-trains, and consequently a momentum-range in these wave-trains in their particle aspect. There is thus an uncertainty δx in the length of the wave-train, and an uncertainty δp in its momentum. We have then

$$\delta x \times \delta p \doteq h.$$

The same uncertainty turns up if we try to observe the total energy E at time t. We then get

$$\delta E \times \delta t \doteq h.$$

Again, if you imagine yourself looking at a sufficiently small particle through a microscope, you can only see it by using a quantum of light which has come from it, and (by the Compton effect) has altered its momentum in doing so. By the elementary theory of optics we know that definition (*i.e.* accuracy of position) varies with the wave-length of light used. The shorter the wave-length used the more accurate the definition (*i.e.* the smaller δx), but the larger the quantum of light received, and hence the larger the alteration in momentum due to the Compton effect of this quantum. Here, if α is the microscope aperture, optical theory gives $\delta x \doteq \dfrac{\lambda}{\sin\alpha}$, and from the Compton effect it can be shown that the component of the momentum-change perpendicular to the axis of the microscope is given by

$$\delta p = \frac{h\nu}{c}\sin\alpha$$

since the Compton recoil is of the order of $\dfrac{h\nu}{c}$.

Thus

$$\begin{aligned}\delta x . \delta p &= \frac{\lambda}{\sin\alpha} \times \frac{h\nu\sin\alpha}{c} \\ &= h \quad (\text{since } \lambda\nu = c).\end{aligned}$$

A very important consequence of both these uncertainties is that, particularly inside a nucleus, one can no longer say (as one should be able to on classical principles, which depended on exact theoretical values of momentum and position) that a particular event is impossible. We can only estimate the (perhaps exceedingly small) probability that it will occur. In particular, if it is a question whether an α-particle will or will not get out of a nucleus past a potential barrier, any indefiniteness in its position or momentum (or energy) must obviously introduce an indefiniteness in the question of whether it will get out; and this brings us to Gamow's theory of the nucleus.

Gamow's Theory

He imagines the situation as illustrated by Fig. 301. From

a long way outside the nucleus the potential rises just as if the inverse square law were obeyed.

At a certain distance from the centre, however, a new law of force becomes effective, and the combination of the two produces a maximum potential at a distance r from the centre. Inside this is a potential "well." PE represents the energy of an α-particle, which according to classical theory could never get over the potential barrier and escape, but according to wave-mechanics (including Heisenberg's principle) has a probability of escape, even if it is a very small one.

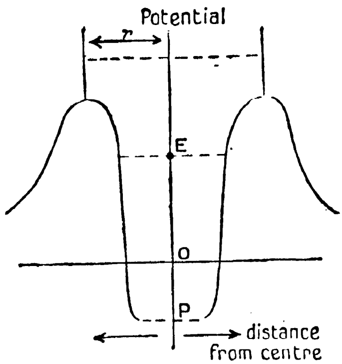


FIG. 301.

For practical purposes we can see that the higher E in the diagram the more frequent is escape to be expected, and the greater the energy of escape. We should thus expect (what we actually find in the Geiger-Nuttall relation) that the shorter the half-life, the smaller the transformation constant λ, the greater the energy of the α-particles and the longer their range.

It is helpful to have a general idea of both Gamow's and Bohr's theories of nuclear structure in one's mind when one is thinking about natural or artificial radioactivity, bombardment, or fission. For example, the possibility, on modern theory, of the escape of an α-particle which would not have enough energy to escape on the classical theory, is balanced by the possibility of penetration by a bombarding α-particle which would not classically have enough energy to penetrate, so that disintegration is possible with particles of much less energy than classical theory would lead us to expect.

There is also the question of penetration, conversion of the nucleus, or fission, by a neutron. The outer Gamow barrier seems to be a coulomb barrier, which would not stop an uncharged particle. Within the well, however, non-coulomb force may be expected to act on an uncharged partsicle like the neutron. Neutrons undoubtedly show resonance. With

certain speeds they can enter certain nuclei, but not with larger or smaller speeds. They seem therefore to be involved in non-coulomb stationary states within the barrier. The Bohr drop theory also gives one a general idea how the introduction of one extra neutron may make the pot boil over and cause fission.

Atomic Fission

The earliest atomic fission seems to have happened, without any one realizing it in 1934 in Rome, where Fermi was trying the recently discovered neutrons on a number of elements, among them uranium. He produced something which showed four separate half-lives, and by chemical methods he showed that the substances causing these activities could not have any atomic number from mercury to uranium, inclusive. He therefore thought he had produced a transuranic element—*i.e.* one with an atomic number greater than 92. This was the situation at the end of 1935. In 1938, following many attempts to identify the mysterious product (including some by Hahn and Strassman (who eventually solved the mystery) in association with Meitner, Curie and Savitch tackled it. They discovered something with a 3·5-hr. half-life. Chemically it could be precipitated with lanthanum as a carrier, and as actinium occupies the same column of the periodic table as lanthanum, they thought it might be a 3·5-hr. isotope of actinium. They also supposed in the end that it must be a new transuranic element, as other evidence was against its being actinium, but they got a step nearer—very warm indeed, in fact—by pointing out that it is difficult to see how an element chemically like lanthanum could occur not too far above uranium to be possible. We now know that they really had got a lanthanum isotope and that if they had grasped this they would have solved the problem of atomic fission first. They had produced fission of U^{238} and the lanthanum was one of the pieces.

Hahn and Strassman then attacked again, and found that if uranium were exposed to neutrons something could be removed chemically, with barium as a carrier which by radioactive decay produced something with the chemical properties of lanthanum. They thought the active cause was really radium, and tried, but failed, to eliminate the barium. It

could not be eliminated, and ultimately the mysterious product was definitely proved to be barium. The uranium atoms had been split by neutrons into large pieces, one of which was barium. Early in January 1939 they announced their discovery, in the interpretation of which they were helped by Meitner and Frisch, to whom the term " fission " is due.

What followed is best described in the words of the account in Pollard and Davidson's *Applied Nuclear Physics* (Chapman & Hall), of the arrival of the news of nuclear fission in America.

" The announcement of this discovery by Hahn and Strassman caused a bee-hive of activity in many places. First word of it in America was received in a telegram to Prof. Niels Bohr, at that time a visiting lecturer at the Institute for Advanced Study, Princeton, New Jersey. Upon learning of this exciting discovery, he made an impromptu address, disclosing it to the many scientists in attendance. The announcement almost broke up the meeting. Every nuclear physicist present realized that, if such a splitting of the uranium atom actually took place under neutron bombardment, the energy release could easily be detected by the simplest kind of research apparatus present in every nuclear laboratory. Consequently those living in and around Washington rushed home to devise an experiment which would allow them to see for themselves. Others burned up the wires to associates back home, urging them to give it a try. The success of these various efforts is to be found in the ' Letters to the Editor ' column of the following issue (15th January 1939) of the *Physical Review*. Six communications concerning fission of uranium appeared, all telling of similar positive results. Under neutron bombardment uranium *did* split into two lighter fragments. Hahn and Strassman were entirely correct." The American experiments provided valuable confirmation of the experimental evidence obtained earlier by Frisch and Meitner.

When the work was continued, it was found that U^{238} was fissioned by fast neutrons, but U^{235} by slow or fast neutrons. As U^{235} also produces the fissioning of neutrons it was at once theoretically possible that in U^{235} a chain reaction can be set up; that is that for every U^{235} nucleus fissioned the resulting neutrons would produce at least one other fission.

This is on the same principle that a race can be maintained so long as on the average every female of one generation produces not less than one female in the next. Without this the race dies. A violently explosive chain reaction in U^{235} depended on two things. First, the U^{235} must be pure. It therefore had to be separated from U^{238}, and U^{235} is only present in a proportion of $\frac{1}{140}$ part of the mixture. The separation problem was immense, and it was solved by the immense industrial resources of America. Secondly, there must be enough uranium in the lump for enough of the escaping neutrons to find a target within the lump. If the lump was too small, too many neutrons escaped and the chain reaction would not go. The way to set the bomb off was thus to get two pieces such that neither of them was large enough to go off by itself, and bring them together suddenly at the right moment into one piece big enough to go off. In the construction of the pile, a problem concerned with the gentle liberation of the energy was that of slowing the neutrons down without absorbing them, and to do this requires substances with special properties. Two of the most useful of these substances are carbon and "heavy" water, which means water in which ordinary hydrogen is replaced by deuterium. This is cheaply produced where water power is available, and no doubt this was why the Germans chose Norway for the heavy water plant which was raided and destroyed in 1942. The existence of this plant was thought to indicate that the Germans were at work on an atomic bomb.

The new transuranic element ${}_{94}Pu^{239}$ was also found fissionable, and it could be produced by bombardment from U^{238}, according to the following reactions:

$${}_{92}U^{238} + n \rightarrow {}_{92}U^{239} + \gamma \text{ is available}$$
$${}_{92}U^{239} \rightarrow {}_{93}Np^{239} + e^- \quad [23 \text{ minutes}]$$
$${}_{93}Np^{239} \rightarrow {}_{94}Pu^{239} + e^- \quad [2{\cdot}3 \text{ days}]$$

A fissionable material is thus produced without the laborious separation of U^{235} from U^{238} (there are other difficulties instead). The Hiroshima bomb was made with U^{235}, the Nagasaki bomb with Pu^{239}. The critical size of the bomb is, of course, secret.

The packing-fraction curve (Fig. 299, p. 638) shows that fission of a heavy element is not the only way to release nuclear energy. Fusion of light elements can equally well do

so, since the lowest packing-fraction energy comes in the middle of the range. The percentage of mass released by fusion is about five times as great as that by fission. For example, a U^{235} nucleus with packing-fraction $(+5{\cdot}2 \times 10^{-4})$ breaks into 2 fragments of average atomic weight 116 and packing-fraction $(-4{\cdot}8 \times 10^{-4})$ and 3 neutrons with mass excess (87×10^{-4}).

$$\text{The mass released} = 10^{-4}[235 \times 5{\cdot}2 - 2 \times 116 \times (-4{\cdot}8) - 3 \times 1 \times 87]$$
$$= 0{\cdot}2075$$

$$\text{Percentage released} = \frac{0{\cdot}2075}{235} \times 100 = {\cdot}088\%$$

If we compare the result of the fusion of 4 tritons, with $_1H^3$ nuclei of mass 3·016997 into 3 helium nuclei, $_2He^4$, of mass 4·00279, we get:

$$\text{Mass reduction} = 4 \times 3{\cdot}01700 - 3 \times 4{\cdot}00279$$
$$= 0{\cdot}05963$$

$$\text{Percentage released} = \frac{0{\cdot}05963}{12{\cdot}06800} \times 100 = 0{\cdot}494$$

about 5·6 times as large a percentage as was released by fission.

The energy of the A-bombs which destroyed Hiroshima (U^{235}) and Nagasaki (Pu^{239}) came from fission. That of the H-bombs is the result of fusion.

A fission-bomb must have a size within definite limits. The Hiroshima bomb released energy equal (we are told) to that of about 20,000 tons of T.N.T. or 2,500 tons of coal of calorific value 8,000 calories per gram. The mass dissipated was therefore given by:

$$\frac{2{,}500 \times 2{,}240 \times 454 \times 8{,}000 \times 4{\cdot}2 \times 10}{9 \times 10^{20}} \text{ grams} = 0{\cdot}9 \text{ grams}$$

so the amount of uranium in the bomb was about

$$0{\cdot}9 \times \frac{100}{0{\cdot}088} \times \frac{1}{454}$$

or between 2 and 3 lb.

Energy from fission can be controlled, and released slowly, and may solve our energy problem when our coal and oil are exhausted.

The H-bomb has no size limits, and until 1955 there seemed to be no way of producing fusion-energy usefully on earth, though it is believed to be the source of the sun's heat.

It seems that our only safeguard against self-destruction by H-bomb is our common sense: a weak reed.

*

QUESTIONS. PART III

These questions are not the type asked in examinations. Their purpose is to enable the reader to test for himself the thoroughness of his reading of the text. When he has got a fair grasp of the text he should be able to answer the questions. Where there is a definite answer it is either given at once or else given in the text.

CHAPTER II

1. Find an expression for the (small) deflection of a particle of charge e and mass m, moving a distance x with velocity v at right angles to an electric field E.

2. Solve question 1 when a magnetic field H is substituted for the electric field E.

3. Show that $\frac{e}{m} = \frac{\theta}{n\mathrm{HT}}$ in Dunnington's method of finding $\frac{e}{m}$

4. Show how to derive Stokes's elementary formula $v = \frac{2}{9}\frac{a^2\rho g}{\eta}$ for the terminal velocity v of a drop of radius a and density ρ falling through a gas of viscosity η.

5. What corrections did Millikan apply to Stokes's formula, and why?

6. In Thomson's positive-ray-parabola experiments show why the curve obtained is a parabola.

7. In Bainbridge's method of finding $\frac{e}{m}$ for positive rays prove that $\frac{e}{m} = \frac{\mathrm{E}}{\mathrm{H_1 H_2} r_1}$.

8. Describe the successive appearances of a simple discharge-tube as it is gradually evacuated. What do you infer from them?

CHAPTER III

1. Show how Planck's equation for E_λ reduces to Wien's for small values of $(\lambda\mathrm{T})$, and to Rayleigh's for large values.

2. On Bohr's first assumption that the electron in its orbit must have an angular momentum which is a multiple of $\frac{h}{2\pi}$, prove

that for circular orbits of radius a in a hydrogen atom $a = \dfrac{n^2h^2}{4\pi^2e^2m}$. Calculate a when $n = 1$, taking $h = 6{\cdot}6 \times 10^{-27}$ erg-secs, $e = 4{\cdot}8 \times 10^{-10}$ e.s.u., $m = 9 \times 10^{-28}$ gm.

3. Accepting Bohr's three assumptions, and taking the nuclear mass as infinite compared with that of the electron, show that

$$\frac{1}{\lambda} = \frac{2\pi^2me^4}{h^3c}\left(\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right).$$

4. From the result of Question 3 calculate R_∞, and compare the result with the value given in the text.

5. Taking $R_\infty = 109{,}737{\cdot}30$ cm.$^{-1}$, and assuming that $m = \dfrac{1}{1837}\,M_H$, calculate R for hydrogen (109,677·58 cm.$^{-1}$).

6. Taking the helium nucleus as $\dfrac{4{\cdot}0028}{1{\cdot}0076}$ as massive as the hydrogen nucleus, find R for singly ionized helium (109,722·26 cm.$^{-1}$).

7. Prove that, for a transition from an orbit of high n to the next orbit below it, the frequency of the line emitted according to Bohr's equation approaches the frequency of the electron in its orbit as $n \to \infty$.

CHAPTER IV

1. Find the energy in ev and the velocity in cm./sec. of an electron giving X-ray wave-length of (a) 1 Å, (b) 4 Å, on impact, with the values of e, m, h used in Question 2, Chapter III, and taking $c = 3 \times 10^{10}$ cm./sec.

[(a) 12,375 volts, $6{\cdot}63 \times 10^9$ cm./sec.
(b) 3,094 volts, $3{\cdot}31 \times 10^9$ cm./sec.).]

2. Prove that if the falling off of intensity I of a beam of X-rays in distance dx is given by $dI = \alpha I dx$, and if I_0 is the intensity when $x = 0$, then $I = I_0 e^{-\alpha x}$.

3. Explain with a rough diagram the principle of Bragg's method of measuring the wave-length of X-rays.

4. Prove that the average life of an atom of a radioactive element is 1·443 of the half-life.

5. Find the velocity of an electron of (a) 3MeV, (b) 11 MeV, and show that the mass of the latter is more than twenty times its rest-mass [Velocities given in text.]

6. Show that 1 ev $= 1{\cdot}6 \times 10^{-12}$ ergs, nearly,
and 1 MeV $= 1{\cdot}6 \times 10^{-6}$ ergs.

7. Explain with diagrams the principle of the G.-M. counter, with special reference to the part of the characteristic curve which is used.

8. Deduce de Broglie's equation $\lambda = \frac{h}{mv}$ on the assumption that the line-integral of the electron's orbit must be a whole number of wave-lengths. (Circular orbits for hydrogen only to be considered.)

9. Deduce from the result of Compton's scattering experiments that if a mass m moving with velocity v is associated with a wave-length at all we may expect that wave-length to be $\frac{h}{mv}$.

10. Show by applying Pauli's principle to the assumption that every electron in an atom is characterized by 4 quantum numbers, that the 1-quantum level should have 2 electrons, the 2-quantum 8, the 3-quantum 18, and the 4-quantum 32.

CHAPTER V

1. Explain the principle of Einstein's photo-electric equation, and show that a γ-ray of wave-length 0·035 Å falling on a surface of work function 70 kilovolts will eject an electron of energy about 275 kilovolts.

2. Give a general idea of the results of the classical theory of the distribution of energy among free electrons in metals, and compare them with the results of the Fermi-Dirac Theory.

3. Define μ, ρ, and g_m for a triode valve, explaining them with reference to the characteristic curves of the valve. Prove that $g_m = \frac{\mu}{\rho}$.

4. Prove that the actual voltage-amplification of a triode with a resistance R in the anode circuit is $\frac{\mu R}{R + \rho}$, and hence that it approaches μ as R approaches infinity.

5. Give a brief account with diagrams of the principles of three types of P.E. cell.

6. Show how to test Richardson's equation by plotting results which give a straight line if the equation is verified, using a diode valve, or a triode with the grid and anode connected together.

7. Explain and contrast briefly, with sketch-diagrams, the principles of the cyclotron, betatron, and synchrotron.

8. Explain the principle of the Van de Graaff accelerator.

CHAPTER VI

1. What observations led to the discovery of the neutron, and why could they not be explained by the theoretical resources then available?

2. What is the evidence that seems to require neutrinos to exist, although they cannot be directly observed? Why does it require their existence?

3. What is the special importance of the deuteron in connection with the development of atomic energy?

4. Why has it been thought that cosmic rays may include more positive than negative primary particles?

5. What is the evidence that cosmic rays originate outside the solar system?

6. What analogy led to Yukawa's 1935 inference about the meson's mass? Why did it lead to a value intermediate between that of the electron and that of the proton?

7. Why did the 1936 observations of Anderson and Neddermayer require a particle intermediate in mass between an electron and a proton?

8. Why is it likely that the meson discovered in 1936 is not the kind of meson predicted by Yukawa in 1935?

9. What is the equation which makes the conception of negative total energy require consideration? And why does this conception exclude negative energies numerically less than m_0c^2?

10. How does this conception lead to an explanation of pair-production, and what is the maximum possible wave-length of a quantum of radiation capable of producing an electron-positron pair? Deduce this value precisely from known numerical values.

11. Sketch the packing-fraction curve for the elements, and explain its meaning. What are the most obvious inferences that can be made from it?

12. Do the same as in Question 11 for the curve of mass excess and defect, and draw inferences.

13. What differences are there in general between what is emitted by natural radioactive sources and what is emitted by artificial radioactive sources?

14. Outline briefly Bohr's drop theory of the nucleus, and give

an indication of how, in a general way, it accounts for natural radioactivity, the emission of particles from stable nuclei under α-particle bombardment, and the fission of heavy atoms after capture of slow neutrons.

15. What is the essential difference between Heisenberg's view of observations in physics, and the classical view as exemplified by Laplace?

16. Show how Heisenberg's Uncertainty Principle is related to Planck's quantum h.

17. Give a brief account of Gamow's theory of the nucleus, and show how it is involved with Heisenberg's Uncertainty Principle.

18. How is it that the Geiger-Nuttall relation, or something of the kind, may be expected on Gamow's theory of the nucleus?

19. Show that if the heat dissipated in an emission of nuclear energy is equivalent to that arising from the combustion of 2,500 tons of coal giving 8,000 calories per gm., then the total mass dissipated is slightly less than 1 gm.

APPENDIX

QUANTUM NUMBERS AND ELECTRONS

The table below shows precisely how the quantum numbers must distribute themselves, according to Pauli's principle (page 557), among the first four "shells" of electrons round an atom. The set of quantum numbers used here is the more usual (but slightly more difficult) set than the n, l, m, σ used on page 558. n is the principal quantum number determining the shell; l the subordinate quantum number determining the sub-shell. When $l = n - 1$ the orbit is circular. The smaller l gets, the more elliptic the orbit. Electrons for which $l = 0$ are s-electrons. Those for which $l = 1$ are p-electrons. Those for which $l = 2$ are d-electrons. Those for which $l = 3$ are f-electrons.

Originally the letters s, p, d, f stood for sharp, principal, diffuse, fundamental, but these words have only historical significance now.

j is $(l + \sigma)$, except that j cannot be negative, so that j is always $\frac{1}{2}$ when $l = 0$, since σ must be $+\frac{1}{2}$ (page 557). j is called the inner quantum number.

m, the magnetic quantum number, differs from the m of page 558 by $\frac{1}{2}$ unit, but has the same effect on the number of electrons satisfying Pauli's principle. m may have any value at integral steps from $+j$ to $-j$. Thus, if $l = 3$, $j = 3\frac{1}{2}$ or $2\frac{1}{2}$. If $j = 2\frac{1}{2}$, m can be $+2\frac{1}{2}, +1\frac{1}{2}, +\frac{1}{2}, -\frac{1}{2}, -1\frac{1}{2}, -2\frac{1}{2}$.

Pauli's principle amounts in practice to the rule that every possible combination of 4 quantum numbers can have one, but only one, electron associated with it. The reader should easily see that the table works out on this principle; and every reader should get sufficiently familiar with the table to be able to compile it without aid on a blank sheet of paper.

A thorough grasp of this table, even without any understanding of wave mechanics, can help the beginner to a most useful understanding of the periodic table, valency, optical

spectra, and *X*-ray spectra, including the cause of multiplet lines in general. For immediate further study by beginners I suggest the following from the list of suggested reference books:

The Periodic Table. No. (5), Born, chap. vi, sec. 6, pp. 169–175.

Valency. No. (23), Speakman, chap. vii, pp. 127–35.

Optical Series Spectra. No. (24), White, chaps. i and ii, and No. (22), Tolansky, chap. viii.

X-ray Series Spectra. No. (5). Born, chap. v, sec. 2, pp. 120–2, which strikes me as particularly terse and clear.

Multiplet Lines. Chap. iii of Tolansky's *Hyperfine Structure in Line Spectra* (Methuen, 6*s*.) gives a very short but clear introduction.

	K (2)		L (8)								M (18)																	
n	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
l	0	0	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
j	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$
m	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{5}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$-\frac{5}{2}$
Sub-group	$1s$		$2s$		$2p$						$3s$		$3p$						$3d$									

	N (32)																														
n	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
l	0	0	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3
j	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{7}{2}$	$\frac{7}{2}$	$\frac{7}{2}$	$\frac{7}{2}$	$\frac{7}{2}$	$\frac{7}{2}$	$\frac{7}{2}$
m	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{5}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$-\frac{5}{2}$	$\frac{5}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$-\frac{5}{2}$	$\frac{7}{2}$	$\frac{5}{2}$	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$-\frac{5}{2}$
Sub-group	$4s$		$4p$						$4d$										$4f$												

USEFUL CONSTANTS

NOTE. When a number in brackets appears at the end of one of the values shown for the constants, the value is quoted from the book referred to by that number in the list of reference books, p. 662. The letter (C) stands for W. H. J. Childs's *Physical Constants* (Methuen). Values without a reference are calculated from the others.

The velocity of light, c,	$= 2{\cdot}99774 \pm 0{\cdot}00010 \times 10^{10}$ cm./sec. (11)
The electronic charge, e,	$= 4{\cdot}8025 \pm 0{\cdot}0004 \times 10^{-10}$ e.s.u. (22)
	$= 1{\cdot}60203 \pm 0{\cdot}00013 \times 10^{-20}$ e.m.u.
Planck's constant, h,	$= 6{\cdot}624 \pm 0{\cdot}002 \times 10^{-27}$ erg-secs. (22)
Avogadro's number, N_0,	$= 6{\cdot}0228 \pm 0{\cdot}0011 \times 10^{23}$ per gm.-molecule (chemical scale). (C)
	$= 6{\cdot}0244 \pm 0{\cdot}0011 \times 10^{23}$ per gm.-molecule (physical scale).
The Faraday, $N_0 e$,	$= 9648{\cdot}7$ e.m.u. per gm.-molecule (chemical scale).
	$= 9651{\cdot}2$ e.m.u. per gm.-molecule (physical scale).
The gas-constant, R_0,	$= 8{\cdot}31436 \pm 0{\cdot}00038 \times 10^{7}$ ergs. per gm.-molecule per degree. (C)
Vol. of 1 gm.-molecule at N.T.P.	$= 22{\cdot}4140 \pm 0{\cdot}0006$ litres. (C)
No. of molecules per c.c. at N.T.P.	$= \dfrac{6{\cdot}0228 \times 10^{23}}{2{\cdot}2414 \times 10^{4}} = 2{\cdot}6871 \times 10^{19}$.
Boltzmann's constant, k,	$= \dfrac{R_0}{N_0} = 1{\cdot}38047 \times 10^{-16}$ ergs. per molecule per degree.
Stefan's constant, σ,	$= 5{\cdot}6728 \pm 0{\cdot}0037 \times 10^{-5}$ ergs/cm.^{2}/sec./deg.^{4}
Rydberg's constant for hydrogen, R_H,	$= 109{,}677{\cdot}58$ cm.$^{-1}$. (2)
Same for helium, R_{He},	$= 109{,}722{\cdot}26$ cm.$^{-1}$. (2)
Same for nucleus of effective infinite mass, R_∞,	$= 109{,}737{\cdot}30 \pm 0{\cdot}017$ cm.$^{-1}$. (2)
$\dfrac{e}{m}$ for electron	$= 1{\cdot}7592 \pm 0{\cdot}0015 \times 10^{7}$ e.m.u. per gm. (11)

Rest-mass of electron, m_0, $= 9{\cdot}1066 \times 10^{-28}$ gm. $\left[= \dfrac{e}{e/m}\right]$.

Mass of H-atom $= 1{\cdot}6748 \pm 0{\cdot}007 \times 10^{-24}$ gm. (11)

Mass of proton $= 1{\cdot}6739 \times 10^{-24}$ gm.

Ratio, $\dfrac{\text{mass of proton}}{\text{rest-mass of electron}}$, $= 1837 = \dfrac{R_\infty - R_H}{R_\infty}$.

Radius a_0 of 1-quantum circular Bohr orbit for hydrogen $= 5{\cdot}27 \times 10^{-8}$ cm. (11)

Mass of 1 mass-unit on chemical scale, on which O $= 16{\cdot}0000$, $= 1{\cdot}66035 \pm 0{\cdot}00031 \times 10^{-24}$ gm. (C)

(This O atom is purely statistical. Its weight is the average weight of an atom of the normal oxygen mixture of isotopes, in which, if we had taken the O^{16} isotope as 16·0000 units, we should have found traces of O^{17} atoms of mass 17·0045 units and O^{18} atoms of mass 18·005 units. The statistical O atom is thus a little heavier than the O^{16} isotope, the ratio of the masses being $\dfrac{16{\cdot}0044}{16{\cdot}0000}$, or 1·000275).

Mass of 1 mass-unit on the physical (or atomic) scale, on which $O^{16} = 16{\cdot}0000$, $= 1{\cdot}65989 \times 10^{-24}$ gm.

The next series is given in atomic mass-units (i.e. on the physical scale.)

Mass of ${}_1H^1$ atom, including electron, $= 1{\cdot}00813$.

Mass of electron $= 0{\cdot}000549$.

Mass of proton $= 1{\cdot}00758$.

Mass of neutron $= 1{\cdot}00893$. (2)

Mass of deuteron, ${}_1H^2$, $= 2{\cdot}01473$. (2)

Mass of triton, ${}_1H^3$, $= 3{\cdot}016997$.

Mass of ${}_2He^4$ atom $= 4{\cdot}00389$. (2)

Mass of α-particle $=$ Mass of ${}_2He^4$ less $2m_0$ $= 4{\cdot}00279$.

1 electron-volt, ev, $= 1{\cdot}60203 \times 10^{-12}$ ergs.

1 MeV, 10^6 ev, $= 1{\cdot}60203 \times 10^{-1}$ ergs.

m_0c^2, the energy associated with m_0, $= 0{\cdot}511$ MeV.

M_1c^2, the energy associated with 1 mass-unit, $= 932$ MeV.

*

METRE-KILOGRAMME-SECOND UNITS

THE M.K.S. system is being adopted by electrical engineers, and physicists should understand it.

It uses the joule, watt, coulomb, volt, ohm, farad, and henry without change of size, and brings all other units into a single system needing no arbitrary constants, at the cost of altering a few of the defining equations and giving the so-called permeability and permittivity of free space extremely peculiar values.

The difficulty of getting a single self-consistent system for all electrical units is like the difficulty of laying a stair-carpet which is a little too long. You can have any stair of your choice flat, but you cannot have every stair flat. There will always be a hump somewhere. The "rationalized" M.K.S. system has two humps, in μ_0, called the "permeability of free space," and ϵ_0, called the "permittivity of free space."

The first new M.K.S. unit that must be understood is the newton, the unit of force.

The newton is the force giving an acceleration of 1 metre per second per second to a mass of 1 kilogramme. It is clearly $10^2 \times 10^3 = 10^5$ dynes, nearly 3·6 oz. weight, or 0·225 lb. weight.

The joule is the M.K.S. unit of work or energy; the work done when a newton moves its point of action through 1 metre. It is therefore $10^2 \times 10^5 = 10^7$ ergs, as usual.

The watt is the M.K.S. unit of power, one joule per second, as usual.

The ampere is the M.K.S. unit of current, but for reasons which will emerge it is defined in a new way. It is now: "the value of a constant current which, when maintained in two parallel rectilinear conductors of infinite length and negligible cross-section, and separated by a distance of 1 metre *in vacuo*, produces between these conductors a force of 2×10^{-7} newtons per metre length."

The M.K.S. unit of magnetizing force is the ampere-turn per metre-length. Whereas for the physicist's oersted $H = 4\pi ni$, with n in turns per cm. and i in e.m. units of 10 amps each, in M.K.S. units $H = nI$ with n in turns per metre and I in amperes.

The field H, in M.K.S. units, at a distance of r metres from a long straight current of I amperes thus becomes $H = \frac{I}{2\pi r}$ instead of suiting the equation of p. 345; and the force F between two long straight currents, I_1 and I_2, r metres apart in free space, becomes

$$F = \frac{\mu_0 I_1 I_2}{2\pi r}$$

per length of 1 metre, not as on p. 345.[1] If we apply the new definition of the ampere to this equation, we get:

$$2 \times 10^{-7} = \frac{\mu_0 \times 1 \times 1\delta}{2\pi \times 1}$$

so that necessarily $\mu_0 = 4\pi \times 10^{-7}$, a value sometimes described as the "permeability of free space," but better called the "magnetic constant of free space."

Since Maxwell's theory gives us:

$$\mu_0 \epsilon_0 = \frac{1}{c^2}$$

where c is the velocity of light, it is necessary that

$$\epsilon_0 = \frac{1}{4\pi c^2 \times 10^{-7}}$$

a quantity sometimes described as the "permittivity of free space." We shall call it for the present the "electric constant of free space."

Since c has the value $2{\cdot}998 \times 10^8$ metres per second, which is very near 3×10^8, it is convenient to put

$$\epsilon_0 = \frac{1}{36\pi \times 10^9}$$

with an error of only about 1 in 1,500.

Direct numerical values often used are:

$$\mu_0 = 1{\cdot}257 \times 10^{-6}$$
$$\epsilon_0 = 8{\cdot}854 \times 10^{-12}$$

A peculiar feature of the M.K.S. system emerges here, for whereas in the c.g.s. systems k and μ apply equally *in vacuo* and for materials (e.g. $k=1$ *in vacuo* and about 6 for glass on the c.g.s. electrostatic system, and $\mu = 1$ *in vacuo*

[1] The reader should be particular to take up the page references in this note. They contain fundamental equations fully discussed in *loc. cit.*, with supporting arguments too long to be repeated in this short introduction to the M.K.S. system.

and about 1,000 for soft iron on the c.g.s. electromagnetic system), on the M.K.S. system we have μ and ϵ with exactly the values they have on the two c.g.s. systems, each to its own. But μ_0 and ϵ_0 remain utterly different from the former μ_0 and k_0.

Thus where in calculating the capacity of a condenser we formerly required only k for the material between the plates, we now need $\epsilon\epsilon_0$; and where in dealing with ferromagnetic materials we formerly needed only μ we now need $\mu\mu_0$. It is as if, whenever we need the permeability or permittivity of material substances, we first use μ_0 or ϵ_0 to get the material into space with the right dimensions, and then use the ratio μ or ϵ to allow for the nature of the material. In short, where the c.g.s. system has k, the M.K.S. system has $\epsilon_0\epsilon$, and where the c.g.s. system has μ, the M.K.S. system has $\mu_0\mu$.[1]

Thus formerly, if H = 10 c.g.s. units and μ max. = 1,600, we had:

$$\text{B max.} = \mu\text{H} = 16{,}000 \text{ c.g.s. units;}$$

now:

$$\begin{aligned}\text{B max.} &= \mu_0\mu\text{H}\\ &= 4\pi \times 10^{-7} \times 1{,}600 \times \frac{10}{0{\cdot}004\pi} \text{ M.K.S. units}\\ &= 1{\cdot}6 \text{ M.K.S. units.}\end{aligned}$$

The system here discussed can be described as the "rationalized symmetrical" M.K.S. system.

Three other M.K.S. systems based on the ampere would have been possible, since any M.K.S. system based on the ampere can be rationalized or unrationalized, symmetrical or unsymmetrical. What then do these words mean?

On p. 32 we see that

$$\text{D} = \frac{k\text{E}}{4\pi} \text{ in the c.g.s. system,}$$

although (p. 358 *passim*) $\text{B} = \mu\text{H}$.

In the rationalized symmetrical M.K.S. system, we have

$$\text{D} = \epsilon_0\text{E and also B} = \mu_0\text{H in free space.}$$

If the factor 4π appears for both B and D or (as in our system) for neither, the system is called symmetrical.

If it appears in neither, the system is called rationalized.

Thus both the c.g.s. systems are unsymmetrical and unrationalized.

[1] But see the note at the end of the text on p. 667, describing another usage that may be met.

One should approach the term "rationalized" with caution. Mr. L. H. A. Carr, in his authoritative paper for the *Metropolitan-Vickers Gazette* of March 1952, says:

"A system is said to be rationalized if the defining equation for each successive unit postulates not only proportionality but also numerical equality between its two sides, except where the properties of free space necessitate the introduction of one of two constants, μ_0 (permeability) and ϵ_0 (permittivity)." The "properties of free space" are, in effect, those which determine the speed of light *in vacuo*.

A complex of π and c has got to crop out somewhere in any system. In other systems it crops out in several places. In the M.K.S. system it is confined to μ_0 and ϵ_0. No mystery is involved except the major mystery of why the constants of nature are what they are.

The other important differences between the M.K.S. and c.g.s. systems constitute definite improvements, for they make for economy of time.

Whereas in the c.g.s. system the electric flux is 4π units from unit charge, in the M.K.S. system it is one unit from one coulomb; so there is never any numerical difference between the magnitude of a charge and the magnitude of the flux from it. We can therefore measure both charge and flux in coulombs, in the M.K.S. system.

In exactly the same way, webers are now the units both for magnetic pole and magnetic flux. As to dimensions, the account given on pp. 471–4 holds if allowance is made for the new equations given on the right in the table below. It will be found that the four ratios given in line 5, p. 474, all become unity. It is good practice to demonstrate this. The modifications give us in effect a new electro-magnetic system based on the ampere and a new electrostatic system based on the coulomb, such that every unit on either system is identical with the corresponding unit on the other system. The M.K.S. system really is two systems adjusted to correspond. It is customary now (1955) not to think as Maxwell thought (p. 471) but to follow the line of argument of p. 473, and say that all we know of the dimensions of ϵ_0 and μ_0 is that $[\epsilon_0\mu_0]=[T]^2[L]^{-2}$, so that $\frac{1}{\sqrt{\epsilon_0\mu_0}}$ is a velocity, the velocity of light *in vacuo*.

A list of M.K.S. units in common use can now be given. It is better to take the system as it stands and use it as a whole without bothering about c.g.s. units, if you use it at all.

Where forces are involved, they cannot help coming out in newtons if the algebra is done correctly.

M.K.S. UNITS

Symbol	*Nature*	*Name*	*New Equation*
I	Current (steady or R.M.S.)	Ampere	
i	Current (instantaneous)	Ampere	
J	Current Density	Amp/(metre)^2	
Q	Charge	Coulomb	
q	Charge (instantaneous)	Coulomb	Force between Charges *in vacuo* $= \frac{QQ^1}{4\pi\epsilon_0 r^2}$ newtons
ψ (psi)	Electric Flux	Coulomb	
D	Electric Flux Density	Coulomb/(metre)^2	
V	Potential Difference or Electromotive Force, (steady or R.M S.)	Volt	$V = \frac{Q}{4\pi\epsilon_0 r}$
v	same (instantaneous)	Volt	
E	Electric Field	Volt/metre	$E = \frac{Q}{4\pi\epsilon_0 r^2}$
ϕ (phi)	Magnetic Flux	Weber	
m	Magnetic Pole	Weber	Force between poles *in vacuo* $= \frac{mm^1}{4\pi\mu_0 r^2}$
B	Magnetic Flux Density	Wb/(metre)^2	
H	Magnetic Field	Ampere-turn/metre	$H = \frac{m}{4\pi\mu_0 r^2}$
F	Magnetomotive Force	Ampere-turn	
S	Reluctance	Ampere-turn/weber	
R	Resistance	Ohm	
Z	Impedance	Ohm	
X	Reactance	Ohm	
ρ (rho)	Resistivity	Ohms per metre cube	
G	Conductance	Mho	
B	Susceptance	Mho	
σ (sigma)	Conductivity	Mhos per metre cube	
L	Self-Inductance	Henry	
M	Mutual Inductance	Henry	
C	Capacitance	Farad	
U	Elastance ($U=1/C$)	Daraf	

		Numerical values of space constants
μ (mu)	Relative Permeability	
μ_0	Permeability of free space, or Magnetic constant of free space	$4\pi/10^7$ $= 1{\cdot}257 \times 10^{-6}$
ϵ (epsilon)	Relative Permittivity	
ϵ_0	Permittivity of free space, or Electric constant of free space	$1/36\pi \times 10^9$ nearly, $= 8{\cdot}854 \times 10^{-12}$

A few other equations should be understood early.

The magnetic potential Ω[1] at a distance r from a pole of strength m is:

$\Omega = \dfrac{m}{4\pi\mu_0 r}$ *in vacuo*, from the defining equation for force between poles, given in the table above,

and hence the magnetic potential of a current I is given by

$\Omega = I\theta$ where θ is the aperture (see pp. 331–2).

The line-integral of a current (see pp. 337–8) is equal to the current, instead of being 4π times the current.

The capacitance of a parallel-plate condenser of area A, d between plates, is given (ignoring end-corrections) by

$$C = \frac{\epsilon\epsilon_0 A}{d} \text{ farads.}$$

It is a surprise to see immediately that point-charges of one coulomb each, one metre apart *in vacuo*, exert a mutual force of

$$\frac{1 \times 1}{4\pi \times \epsilon_0 \times 1} \text{ newtons} = 9 \times 10^9 \text{ newtons} = 900{,}000 \text{ tons weight.}$$

NOTE. It has also been recommended that instead of making ϵ and μ the relative permittivity and permeability of a material, these symbols should be again, as in c.g.s. systems, the absolute permittivity and permeability. The relative permittivity and permeability then become ϵ_r and μ_r, so that $\epsilon_r = \dfrac{\epsilon}{\epsilon_0}$, $\mu_r = \dfrac{\mu}{\mu_0}$. With this usage, ϵ_r and μ_r are pure numbers.

[1] The symbol Ω is a capital form of the Greek letter ω, used for electrical resistance and for its unit, the ohm. Ω and ω are alike pronounced "omega."

REFERENCE BOOKS FOR ATOMIC PHYSICS

THIS note refers to books which might be helpful for physics students in their last two years at school and their first two at the university. I feel that in present conditions, when one must do the best one can with the books one can find and afford, it is better to give a long list in the hope that in each section at least one may be easily accessible. The county libraries can and will do an astonishing amount to get books of this type, freely posted and often from distant libraries, at the request of individual students.

Dates of publication are not given. It may be assumed that a reader should always go for the latest edition he can get hold of.

1. I. Tolansky: *Introduction to Atomic Physics*, Longmans.
2. F. K. Richtmyer and E. H. Kennard: *An Introduction to Modern Physics*, McGraw-Hill.
3. J. D. Stranathan: *The 'Particles' of Modern Physics*, The Blakiston Company, Philadelphia.
4. E. Pollard and W. L. Davidson: *Applied Nuclear Physics*, John Wiley & Sons, New York; Chapman & Hall, London.

Nos. 2, 3, and 4 are written with the astonishing gusto one finds among American physicists to-day. Each of them told me plenty of things I did not find in the other two, specially about the periodic table, atomic weights, isotopes, and new radioactive substances. A sound average physicist could understand practically every word of 3 and 4, but parts of 2 would beat him. 2 has a remarkable historical introduction, getting from Thales of Miletus (600 B.C.) to Maxwell's equations of the electromagnetic field in seventy pages. 3 has the most details and the most references to original papers of any physics text-book I have seen. 4 is on a smaller scale and covers less general ground (as its title implies). The authors write like a witty man talking, and it is difficult to stop reading the book whenever one picks it up. 3 has the biggest collection of numerical questions about atomic physics that I

know of. If I could only have one book on atomic physics this would be it.

5. Max Born, F.R.S.: *Atomic Physics*, Blackie & Son.

This book is by the Tait Professor of Natural Philosophy at the University of Edinburgh, and one of the three originators, with Heisenberg and Jordan, of quantum mechanics. The book is really for Final Honours or Post-graduate students. It has in many places the feeling of giving one direct the original thoughts of one of the men who are really at the centre of the work. Most of it is too advanced for our purposes, but even for an intermediate student it is most exhilarating to look into. Very often a problem about which one is hazy is suddenly jerked into focus in a few words. Any one who intends to be a real theoretical physicist should have it handy.

6. J. B. Hoag: *Electron Physics*, Van Nostrand, New York.
7. G. Friedlander and J. W. Kennedy: *Introduction to Radio-chemistry*, Chapman & Hall.

Both 6 and 7 are practical handbooks.

8. J. J. Thomson: *Rays of Positive Electricity*, Longmans.
9. F. W. Aston: *Isotopes*, Arnold.
10. J. Perrin: *Atoms*, Constable.
11. R. A. Millikan: *Electrons*, University of Chicago Press.

These last four are source books of original work, and have the special clearness of most accounts of work described by the man who did it.

12. J. W. Cork: *Radioactivity and Nuclear Physics*, Van Nostrand, New York.
13. A. H. Compton and S. K. Allison: *X-Rays*, Macmillan.
14. J. K. Roberts: *Heat and Thermodynamics*, Blackie.
15. L. de Broglie: *The Revolution in Physics*, Routledge.
16. Max Born: *The Restless Universe*, Blackie.
17. C. G. Darwin: *The New Conceptions of Matter*, Bell.
18. E. N. da C. Andrade: *The Structure of the Atom*, Bell.
19. H. E. White: *Introduction to Atomic Spectra*, McGraw-Hill.

This is a most satisfying standard text-book beginning with a long and clear introductory historical chapter, which specially helps the reader to understand how the various notations for spectra originated and developed.

I do not feel that a student should make up his mind to read the whole of any of the above books except 3 and 4. These two, though packed with matter, are almost as easy to read as a good Agatha Christie detective story, evidently because the authors so much enjoyed writing them. Professor Born must have enjoyed writing 5 and 16 just as much, but 5 contains some concentrated difficult theory and 16 is a 'popular' book covering a relatively narrow range.

As a student I generally found my mind possessed by some unsolved problem, and I always felt that there was somewhere an author who looked at it in a way which suited me. I needed plenty of books to help me find the key author for each problem. I did not have to look long in a book to see if it was my sort or not. For example, 15, 16, 17, 20, 21, 22, 23 discuss rather similar problems in very different ways. Nos. 21, 22, and 23 are in effect one big book, and for myself this big book was the clearest. All these seven books seem to me good on their difficult subject, but (for any particular student) one well studied may be enough.

20. F. A. Lindemann: *Quantum Theory*, Oxford.

21. A. S. Eddington: *The Nature of the Physical World*, Cambridge.

22. A. S. Eddington: *The New Pathways in Science*, Cambridge.

23. A. S. Eddington: *The Philosophy of Physical Science*, Cambridge.

24. G. T. P. Tarrant: *Electricity, Magnetism, and Modern Physics*, Dent.

25. G. R. Noakes: *Text-book of Electricity and Magnetism*, Macmillan.

26. M. Nelkon: *Physics and Radio*, Arnold.

27. E. Williams: *Thermionic Valve Circuits*, Pitman.

28. G. Gamow: *Birth and Death of the Sun*, Macmillan.

29. G. Gamow: *Creation of the Universe*, Viking Press.

30. G. Gamow: *Biography of the Earth*, Macmillan.

31. Colston Society, Bristol: *Cosmic Radiation*, Butterworth.

32. D. J. X. Montgomery: *Cosmic Ray Physics*, Oxford.

33. R. E. Peierls: *Laws of Nature*, Allen & Unwin.

34. W. F. Magie: *Source Book in Physics*, McGraw-Hill.

Methuen's Monographs on Physical Subjects are very comprehensive. Each book is small enough for the pocket. *Review of Modern Physics*, published four times a year by the American Institute of Physics, Lancaster, Pennsylvania, contains (in addition to much difficult mathematics for specialists) occasional clear and comprehensive accounts of special subjects.

All students of physics should look through every number of the Penguin Science News Series, as it comes out. Every number has an up-to-date Research Report, and most have very clear and good articles on special subjects in atomic physics. These articles are unlike anything else now available. Though authoritative and full of numerical details and clear diagrams, they are all extremely clear and easy to read. No. 8 came out in the summer of 1948, No. 37 in August 1955. Many of the articles are by young research workers.

INDEX